

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 10:38:48 ; Search time 433.634 Seconds
(without alignments)
15657.843 Million cell updates/sec

Title: us-09-762-767a-1
Perfect score: 3015
Sequence: 1 cccttgcgcgattggcgg.....aaaaaaaaaaaaaaaaaaaaa 3015

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3015	100.0	3015	21	Human soluble guan
2	2950.8	97.9	4742	23	Human prostate exp
3	2950.8	97.9	4742	23	Human prostate exp
4	2950.8	97.9	4742	23	Human prostate exp
5	2950.8	97.9	4742	23	Human prostate exp
6	2855.4	94.7	3004	24	Prostate cancer-as
7	2855.4	94.7	3004	24	Prostate cancer-as
8	2368	78.5	2715	21	Human cyclic nucle
9	745.8	24.7	885	22	Human neuroblastom

10	548	18.2	2954	24	AAS94827	Human DNA sequence
11	482.2	16.0	761	22	AA23886	Human breast cance
12	315.6	10.5	2357	23	ABL02451	Drosophila melanog
13	314.8	10.4	5259	23	ABL02450	Drosophila melanog
14	308.4	10.2	649	22	ABL15033	Human breast cance
15	260.8	8.7	3094	23	ABL04726	Drosophila melanog
16	204.4	6.8	751	22	AA197899	Human neuroblastom
17	192.6	6.4	2443	21	AA288939	Human soluble guan
18	192.6	6.4	3196	22	AAH98392	Human EST-derived
19	176.4	5.9	2335	24	ABK63662	Rat sequence diffe
20	155.6	5.2	2850	23	ABL02665	Drosophila melanog
21	152.4	5.1	5424	24	ABL02854	Human immune syste
22	149.4	5.0	3821	22	ABA09027	Human ANP-A recept
23	148.6	4.9	3519	23	ABL30435	Drosophila melanog
24	148.6	4.9	6756	23	ABL30434	Drosophila melanog
25	147	4.9	2212	23	ABL07087	Drosophila melanog
26	146	4.8	3697	12	AAQ10324	Human Natriuretic
27	144.4	4.8	4081	24	ABR83833	Human cDNA diffe
28	139.8	4.6	3231	23	ABL05735	Drosophila melanog
29	139.8	4.6	6339	23	ABL05734	Drosophila melanog
30	139.6	4.6	2928	23	ABL15689	Drosophila melanog
31	135	4.5	3286	23	ABL28779	Drosophila melanog
32	134	4.4	8895	23	ABL28778	Drosophila melanog
33	128.8	4.3	2766	23	ABL29938	Drosophila melanog
34	128.2	4.3	3762	22	AAF89876	Nucleotide sequenc
35	127.4	4.2	5565	23	ABL29715	Drosophila melanog
36	127.4	4.2	5424	24	ABL32855	Human immune syste
37	125.6	4.2	3784	14	AAQ47356	GC-C DNA. Rattus
38	124.6	4.1	3499	22	ABF89878	Nucleotide sequenc
39	121.4	4.0	2664	23	AA567254	DNA encoding novel
40	121.4	4.0	2664	23	AA583974	DNA encoding novel
41	121.4	4.0	2664	23	AA589581	DNA encoding novel
42	121.4	4.0	3786	18	AA789300	Human ST receptor
43	121.4	4.0	3786	19	AA797229	Human ST receptor
44	121.4	4.0	3787	24	ABL57068	Human ST receptor
45	121.4	4.0	3787	24	AA516116	DNA encoding human

ALIGNMENTS

RESULT 1
AAZ88938
ID AAZ88938 standard; DNA; 3015 BP.

XX AC AAZ88938;

XX DT 26-MAY-2000 (first entry)

XX DE Human soluble guanylylcyclase alpha DNA.

XX DE Human; guanylylcyclase alpha; hsgCalphal; hsgCBeta1; soluble;
KW guanylylcyclase beta1; antiarteriosclerotic; vasotropic; hypotensive;
KW gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;
KW peripheral arterial occlusive disease; arterial hypertension; ds.

XX OS Homo sapiens.

XX PN DE19837015-Al.

XX PD 24-FEB-2000.

XX PF 14-AUG-1998; 98DE-1037015.

XX PR 14-AUG-1998; 98DE-1037015.

XX PA (VASO-) VASOPHARM BIOFTECH GMBH & CO KG.

XX PI Schmidt H, Zabel U, Poller W;

XX DR WPI; 2000-184044/17.

XX P-PSDB; AA51607.

PT New human soluble guanylate cyclase alpha/beta1 and the nucleic acid
 PT encoding the subunits, useful for producing diagnostic antibodies, and
 PT for somatic gene therapy of arteriosclerosis -

XX Example 1; Page 12-13; 44pp; German.

CC This invention describes novel purified human soluble guanylate cyclase
 CC alpha/beta1 (hsGcalpha/beta1). The products of the invention have
 CC antiarteriosclerotic, vasotropic and hypotensive activity. Nucleic acid
 CC sequences encoding the alpha1 and/or beta1 subunit are useful for somatic
 CC gene therapy of arteriosclerosis and restenosis, ischemia (infarct),
 CC peripheral arterial occlusive disease and arterial hypertension.
 CC Antibodies to hsGcalpha/beta1 can be used for diagnosis of aberrant
 CC hsGcalpha/beta1 expression in human tissues. This sequence encodes the
 CC human soluble guanylylcyclase alpha1 subunit described in the method
 CC of the invention.

XX Sequence 3015 BP; 877 A; 580 C; 716 G; 742 T; 0 other;

Query Match 100.0%; Score 3015; DB 21; Length 3015;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCCTTATGGCGATGGCGGCTGCAGAGACCAGGACTCAGTTCCCTGCGCTAGTCTGAG	60
DB	1	CCCTTATGGCGATGGCGGCTGCAGAGACCAGGACTCAGTTCCCTGCGCTAGTCTGAG	60
QY	61	CCCTAGTGGTGGGACTCAGCTCAGAGTCAGTTTCCAGAGCAGGTTTCCAGTGCAGATT	120
DB	61	CCCTAGTGGTGGGACTCAGCTCAGAGTCAGTTTCCAGAGCAGGTTTCCAGTGCAGATT	120
QY	121	TTCCCTACACTTTTCCCTGCGCTAGAGCAGGAGCAGCTTGGAGCAGACCAGCGGAGGAC	180
DB	121	TTCCCTACACTTTTCCCTGCGCTAGAGCAGGAGCAGCTTGGAGCAGACCAGCGGAGGAC	180
QY	181	ACCTGTGGGGAGGAGCGCTGAGAGGCTTACAGACCCAGCCGCGGGTGATCTCACC	240
DB	181	ACCTGTGGGGAGGAGCGCTGAGAGGCTTACAGACCCAGCCGCGGGTGATCTCACC	240
QY	241	ATGTGCGGATTTGCGAGGCGCCCTGGAGCTGCTAGAGATCCGGAAGCAGACGCCCGGAG	300
DB	241	ATGTGCGGATTTGCGAGGCGCCCTGGAGCTGCTAGAGATCCGGAAGCAGACGCCCGGAG	300
QY	301	GTGTGCGAAGCCACCAAGACTGCGGCTCTTGGAGAAAGGTGAGAGGGGGCCACCGCGG	360
DB	301	GTGTGCGAAGCCACCAAGACTGCGGCTCTTGGAGAAAGGTGAGAGGGGGCCACCGCGG	360
QY	361	TCCTCCGGCTGTCTGCACCCCTGCGCTGAGCTGCTGACAGTGACATGACATCCAC	420
DB	361	TCCTCCGGCTGTCTGCACCCCTGCGCTGAGCTGCTGACAGTGACATGACATCCAC	420
QY	421	TTACCAAGTGTCTTGAATGATAGTGGCTTCTGTTTGTGAGTCTCATATAAGAACTACAG	480
DB	421	TTACCAAGTGTCTTGAATGATAGTGGCTTCTGTTTGTGAGTCTCATATAAGAACTACAG	480
QY	481	CTCATCAGGAGGAGATCGCAGAGGTGAGAGACACCAACCATCTTCTGACGAGCT	540
DB	481	CTCATCAGGAGGAGATCGCAGAGGTGAGAGACACCAACCATCTTCTGACGAGCT	540
QY	541	CAAGGATCTCAAGATCACAGGAGGTGCTCTTCTTCTTACTGGCACCAGGTCAAGTCC	600
DB	541	CAAGGATCTCAAGATCACAGGAGGTGCTCTTCTTCTTACTGGCACCAGGTCAAGTCC	600
QY	601	TAAACGAGTCTTACAGAGGAGCAGCAGAGGCTCAGAGAGCTGCAAGAACCCGTGCCAT	660
DB	601	TAAACGAGTCTTACAGAGGAGCAGCAGAGGCTCAGAGAGCTGCAAGAACCCGTGCCAT	660
QY	661	CTGTCAAGACATTTCTTGAAGAAACATACAGAAAGTCTTCTTCAAGAAACACAGTCG	720
DB	661	CTGTCAAGACATTTCTTGAAGAAACATACAGAAAGTCTTCTTCAAGAAACACAGTCG	720
QY	721	GAGCCGAGTCTATCTCACACTTTGGCAGAGAGTATTTGCAAACTGATTTTCCAGATT	780
DB			

DB	721	GAGCCGAGTCTATCTTACACTTTGGCAGAGAGTATTTGCAAACTGATTTTCCCAAGATT	780
QY	781	TGAACGGCTGAATGTTGCATCTCAGAGAACTTGGCAAGCACAATAAAGAAAGACAG	840
DB	781	TGAACGGCTGAATGTTGCATCTCAGAGAACTTGGCAAGCACAATAAAGAAAGACAG	840
QY	841	GAATCTTTGGAAAGAGAGACTTTGAAAAACAATTCGAGAGCAGAGTTGACAGG	900
DB	841	GAATCTTTGGAAAGAGAGACTTTGAAAAACAATTCGAGAGCAGAGTTGACAGG	900
QY	901	AGTTCAGTGGAGGTATCAAGAATCTCTGGTGAAGAGTTTAAATATGTACGA	960
DB	901	AGTTCAGTGGAGGTATCAAGAATCTCTGGTGAAGAGTTTAAATATGTACGA	960
QY	961	GGAAGTGAATAACATCTTGGGTGGTGGAGGACCTTAAAGATTTTAAACAGCTT	1020
DB	961	GGAAGTGAATAACATCTTGGGTGGTGGAGGACCTTAAAGATTTTAAACAGCTT	1020
QY	1021	CAGTACCCTTCTGAAACAGAGCAGCCATTTGCCAAGACAGCAAAAGGGCAGGCTTGA	1080
DB	1021	CAGTACCCTTCTGAAACAGAGCAGCCATTTGCCAAGACAGCAAAAGGGCAGGCTTGA	1080
QY	1081	GGACGCTCCATCTCTATGCTGGATAGGAGGATGATTTCTACATGTTTACTACTTCTT	1140
DB	1081	GGACGCTCCATCTCTATGCTGGATAGGAGGATGATTTCTACATGTTTACTACTTCTT	1140
QY	1141	CCCTAAGAGAACCACTCCCTGATCTTCCGGCATCAATAAGGCAGCTGCTCAGTATT	1200
DB	1141	CCCTAAGAGAACCACTCCCTGATCTTCCGGCATCAATAAGGCAGCTGCTCAGTATT	1200
QY	1201	ATATGAAGAGGAGTGGAGTGTCTTAATGCTCCCTGCTTCCATATGATGCGAGGA	1260
DB	1201	ATATGAAGAGGAGTGGAGTGTCTTAATGCTCCCTGCTTCCATATGATGCGAGGA	1260
QY	1261	GTTCGTGATCAGCCCTACTTGTGCTACTCGTTCACATGAAGACCAAGCCATCCCT	1320
DB	1261	GTTCGTGATCAGCCCTACTTGTGCTACTCGTTCACATGAAGACCAAGCCATCCCT	1320
QY	1321	GTCCCCAGCAAAACCCAGTCTCTGCTGCTGATTTCCACATCGCTATTTCTGCAAGACAT	1380
DB	1321	GTCCCCAGCAAAACCCAGTCTCTGCTGCTGATTTCCACATCGCTATTTCTGCAAGACAT	1380
QY	1381	TCCATTCATTTTCATGTTTGACAAAGATATGACAAATTTCTGCAATTTGCAATGGATCAG	1440
DB	1381	TCCATTCATTTTCATGTTTGACAAAGATATGACAAATTTCTGCAATTTGCAATGGATCAG	1440
QY	1441	AAGGCTGATGAACAGGAGAGACTTCAAGGAAAGCTTAATTTTGAAGAAATACTTTGAAAT	1500
DB	1441	AAGGCTGATGAACAGGAGAGACTTCAAGGAAAGCTTAATTTTGAAGAAATACTTTGAAAT	1500
QY	1501	TCTGACTCCAAAAATCAACAGACGTTTACGGGATCATGACTATGTTGAAATGACGTT	1560
DB	1501	TCTGACTCCAAAAATCAACAGACGTTTACGGGATCATGACTATGTTGAAATGACGTT	1560
QY	1561	TGTTGTAGAGTGGAGATGGGACACTCTGTGAAGAAATCTTCAAGGGTTTATGGACCT	1620
DB	1561	TGTTGTAGAGTGGAGATGGGACACTCTGTGAAGAAATCTTCAAGGGTTTATGGACCT	1620
QY	1621	CAAGGCGCAATGATCTACATTTGTAATCCAGTCAATCTTGTGTTTGGGTGACCCCTG	1680
DB	1621	CAAGGCGCAATGATCTACATTTGTAATCCAGTCAATCTTGTGTTTGGGTGACCCCTG	1680
QY	1681	TGTGACAGATTTAGAAGATTTTACAGGACGAGGCTCTACCTCTCAGACATCCCAATTCA	1740
DB	1681	TGTGACAGATTTAGAAGATTTTACAGGACGAGGCTCTACCTCTCAGACATCCCAATTCA	1740
QY	1741	CAATGCATGAGGATGTGGTCTTAATAGGGGAACAGCCCGAGCTCAAGATGGCCTGAA	1800
DB	1741	CAATGCATGAGGATGTGGTCTTAATAGGGGAACAGCCCGAGCTCAAGATGGCCTGAA	1800
QY	1801	GAAGAGGCTGGGAAAGCTGAAGGCTTACCTTGAAGCCCAAGCCCTGGAGGAGGA	1860
DB	1801	GAAGAGGCTGGGAAAGCTGAAGGCTTACCTTGAAGCCCAAGCCCTGGAGGAGGA	1860

QY 1861 GAAGAAAAGACAGTAGACCTTCTGTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAGCT 1920
DB |||||||
QY 1861 GAAGAAAAGACAGTAGACCTTCTGTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAGCT 1920
DB |||||||
QY 1921 GTGGCAAGGGAAGTTGTGCAAGCCCAAGATTCAGTAATGTCACCATGCTCTTCTCAGA 1980
DB |||||||
QY 1921 GTGGCAAGGGAAGTTGTGCAAGCCCAAGATTCAGTAATGTCACCATGCTCTTCTCAGA 1980
DB |||||||
QY 1981 CATCGTTGGGTTCACTGCCATCTGCTCCAGTGTCTACCGCTGCAGGTCAACCATGCT 2040
DB |||||||
QY 2041 CAATGCACGTACACTCCCTCGACAGCAGTGTGGAGCTGTGATGTCTCAAGGTGGA 2100
DB |||||||
QY 2041 CAATGCACGTACACTCCCTCGACAGCAGTGTGGAGCTGTGATGTCTCAAGGTGGA 2100
DB |||||||
QY 2101 GACCAATGGCGATGCTATTTGTGTAGCTGGGGATTCACAAAGAGAGTGTACTCATGTC 2160
DB |||||||
QY 2101 GACCAATGGCGATGCTATTTGTGTAGCTGGGGATTCACAAAGAGAGTGTACTCATGTC 2160
DB |||||||
QY 2161 TGTTCCAGATAGCGCTGTAGGCTGAGGATGATGAGCTCTCTGATGAAGTTATGCTCC 2220
DB |||||||
QY 2161 TGTTCCAGATAGCGCTGTAGGCTGAGGATGATGAGCTCTCTGATGAAGTTATGCTCC 2220
DB |||||||
QY 2221 CCATGGAGAACCTATCAAGATGCGAATTTGGAGTGCACCTCTGGATCGATTTTGTGGGCT 2280
DB |||||||
QY 2221 CCATGGAGAACCTATCAAGATGCGAATTTGGAGTGCACCTCTGGATCGATTTTGTGGGCT 2280
DB |||||||
QY 2281 CGTTGGAGTTAAATGCCCCCTTACTGCTTTTGGAAACATGTCACCTGCTGGCTAAACAA 2340
DB |||||||
QY 2281 CGTTGGAGTTAAATGCCCCCTTACTGCTTTTGGAAACATGTCACCTGCTGGCTAAACAA 2340
DB |||||||
QY 2341 ATTTGAGTCCCTGAGTACCAAGAAATCAATGTCAGCCCAACCACTTACAGATTACT 2400
DB |||||||
QY 2341 ATTTGAGTCCCTGAGTACCAAGAAATCAATGTCAGCCCAACCACTTACAGATTACT 2400
DB |||||||
QY 2401 CAAGACACTGCTGTTGCTGTTTACCCCTCGATCAAGGAGGAACTTCCACCAAACTT 2460
DB |||||||
QY 2401 CAAGACACTGCTGTTGCTGTTTACCCCTCGATCAAGGAGGAACTTCCACCAAACTT 2460
DB |||||||
QY 2461 CCTAGTGAATATCCCGGAATCTGCAATTTCTGGATGCTTACCAACAGGAAACAACTC 2520
DB |||||||
QY 2461 CCTAGTGAATATCCCGGAATCTGCAATTTCTGGATGCTTACCAACAGGAAACAACTC 2520
DB |||||||
QY 2521 AAACCATGCTTCCAAAAGAAAGATGTGGAAGATGCAATGCCAATTTTATAGGCAAAAGC 2580
DB |||||||
QY 2521 AAACCATGCTTCCAAAAGAAAGATGTGGAAGATGCAATGCCAATTTTATAGGCAAAAGC 2580
DB |||||||
QY 2581 ATCAGGAATAGATTAGCAACCTATATACCTATTTATAGTCTTTGGGTTGACTCATTTG 2640
DB |||||||
QY 2581 ATCAGGAATAGATTAGCAACCTATATACCTATTTATAGTCTTTGGGTTGACTCATTTG 2640
DB |||||||
QY 2641 AAGATGTGTAGAGCTCTGAAGACACTTTAGGATGTAGATGGCTTAACAAGCAGTATTA 2700
DB |||||||
QY 2641 AAGATGTGTAGAGCTCTGAAGACACTTTAGGATGTAGATGGCTTAACAAGCAGTATTA 2700
DB |||||||
QY 2701 AAATTTAGGAGCCCAAGTCACAACTTTCTCTCTGTTTAACTGACAAAATGACTCATTT 2760
DB |||||||
QY 2701 AAATTTAGGAGCCCAAGTCACAACTTTCTCTCTGTTTAACTGACAAAATGACTCATTT 2760
DB |||||||
QY 2761 CAGTACTTCAGCTCTTCAGAAAACCACTTTAAAGACACTTTTAAAGCTTCTTTGGGAGTA 2820
DB |||||||
QY 2761 CAGTACTTCAGCTCTTCAGAAAACCACTTTAAAGACACTTTTAAAGCTTCTTTGGGAGTA 2820
DB |||||||
QY 2821 TTTCTATTATATAACCAAGCACTTACTACCTGTACTCAAAATTCAGCACCTTGTACATATA 2880
DB |||||||
QY 2881 TCAGATAATTTAGTCAATTTGACAACTGTATGGAGTACACCTGCAATCTCATATCTCTGCT 2940
DB |||||||

QY 2941 GGAATGCCATGGTTATTAAAGTGTGTTTGTGATAGTGTGCTCAAAAAA 3000
DB |||||||
QY 2941 GGAATGCCATGGTTATTAAAGTGTGTTTGTGATAGTGTGCTCAAAAAA 3000
DB |||||||
QY 3001 AAAAAA 3015
DB |||||||
QY 3001 AAAAAA 3015
DB |||||||
RESULT 2
ABV21348
ID ABV21348 standard; cDNA; 4742 BP.
XX ABV21348;
AC ABV21348;
XX
XX
DT 13-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 21339.
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
KW Homo sapiens.
XX
XX WO200160860-A2.
PN
PD 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
PF
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Schlegel R, Endege WO, Monahan JE;
PI
XX WPI; 2001-662795/76.
XX
DR Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 3547; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 4742 BP; 1450 A; 933 C; 998 G; 1355 T; 6 other;
XX
Query Match 97.9%; Score 2950.8; DB 23; Length 4742;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2978; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
QY 6 ATGGCGATTGGCGGCTGCAGAGACAGGACATCCCTCCCTAGCTCTGAGCCTAG 65
|||||

Qy	426	AGTGTCTCTTGAATTGATAGTGGCTTCCTGTTTGTCCAGTCTCATATAAGAACTTACAGCTCAT	485
Db	425	AGTGTCTCTTGAATTGATAGTGGCTTCCTGTTTGTCCAGTCTCATATAAGAACTTACAGCTCAT	484
Qy	486	CAGGAGAGATCGCAGCAGGGTAGAGACACCAACACCATGTTCTGCACGAGCTCAGG	545
Db	485	CAGGAGAGATCGCAGCAGGGTAGAGACACCAACACCATGTTCTGCACGAGCTCAGG	544
Qy	546	ATCTCAAGATCACAGGAGTGTCTTCTTCTCTTACTGGCACCAAGGTCAAGTCTCCTAACG	605
Db	545	ATCTCAAGATCACAGGAGTGTCTTCTTCTCTTACTGGCACCAAGGTCAAGTCTCCTAACG	604
Qy	606	AGTCTTCAGAGGAGCGACGAAAGCTCAGAGAGCTGCAAAAGCAACCGTGCCCATCTGTC	665
Db	605	AGTCTTCAGAGGAGCGACGAAAGCTCAGAGAGCTGCAAAAGCAACCGTGCCCATCTGTC	664
Qy	666	AAGACATCTCTGAGAAGAACATCAAGAAAGTCTTCTCTCAAGGAAAACCAAGTCGGAGCC	725
Db	665	AAGACATCTCTGAGAAGAACATCAAGAAAGTCTTCTCTCAAGGAAAACCAAGTCGGAGCC	724
Qy	726	GAGTCTATCTTCACACTTTGGCAGAGAGTATTTGCAAACTGATTTTCCCAGAGTTTGAAC	785
Db	725	GAGTCTATCTTCACACTTTGGCAGAGAGTATTTGCAAACTGATTTTCCCAGAGTTTGAAC	784
Qy	786	GGCTGAATGTTGCATCTCAGAGAACATTTGGCAAGACACAAATATAAGAAAGCAGGAAT	845
Db	785	GGCTGAATGTTGCATCTCAGAGAACATTTGGCAAGACACAAATATAAGAAAGCAGGAAT	844
Qy	846	CTTTGGAAAAGAGAGACTTTTGA AAAACAATTTGCAGAGCAAGCAAGTTGCAGCAGAGTTC	905
Db	845	CTTTGGAAAAGAGAGACTTTTGA AAAACAATTTGCAGAGCAAGCAAGTTGCAGCAGAGTTC	904
Qy	906	CAGTGGAGGTTATCAAAAGATCTCTTGSTGAAGAGGTTTTTAAATATGTTACAGGAAG	965
Db	905	CAGTGGAGGTTATCAAAAGATCTCTTGSTGAAGAGGTTTTTAAATATGTTACAGGAAG	964
Qy	966	ATGAAACATCCTTGGGGTGGTGGAGGCACCTTAAAGATTTTTTAAACAGCTTCAGTA	1025
Db	965	ATGAAACATCCTTGGGGTGGTGGAGGCACCTTAAAGATTTTTTAAACAGCTTCAGTA	1024
Qy	1026	CCCTTCTGAAACAGAGCAGCCATTTGCCAAGAGCAGGAAAAGGGCAGGCTTCAGGACG	1085
Db	1025	CCCTTCTGAAACAGAGCAGCCATTTGCCAAGAGCAGGAAAAGGGCAGGCTTCAGGACG	1084
Qy	1086	CCTCCATTCATGCTTGGATGAAGAGGATGATTTTCTACATGTTTACTACTTCTTCCTTA	1145
Db	1085	CCTCCATTCATGCTTGGATGAAGAGGATGATTTTCTACATGTTTACTACTTCTTCCTTA	1144
Qy	1146	AGAGAACCCCTCCCTGATTTCTCCCGGCATCATAAAGGACGCTGCTCAGTATTATATG	1205
Db	1145	AGAGAACCCCTCCCTGATTTCTCCCGGCATCATAAAGGACGCTGCTCAGTATTATATG	1204
Qy	1206	AAACGGAAGTGGAGTGTCTGTTAATGCTCCCTGCTTCCATTAATGATTGACGAGGTTTG	1265
Db	1205	AAACGGAAGTGGAGTGTCTGTTAATGCTCCCTGCTTCCATTAATGATTGACGAGGTTTG	1264
Qy	1266	TGAATCAGCCCTACTTGTGTGATCTCCGTTTCACATGAAAGACCAAGCCATCCCTGTCCC	1325
Db	1265	TGAATCAGCCCTACTTGTGTGATCTCCGTTTCACATGAAAGACCAAGCCATCCCTGTCCC	1324
Qy	1326	CCAGCAACCCAGTCCCTCGTGTGATTTCCACATCGCTATTCTGCAAGACATTTCCAT	1385
Db	1325	CCAGCAACCCAGTCCCTCGTGTGATTTCCACATCGCTATTCTGCAAGACATTTCCAT	1384
Qy	1386	TCCATTTTCATGTTGAAAAAGATGACAAATTTGCAATTTGGCAATGGCATCAGAAGGC	1445
Db	1385	TCCATTTTCATGTTGAAAAAGATGACAAATTTGCAATTTGGCAATGGCATCAGAAGGC	1444
Qy	1446	TGATGAACAGAGAGACTTTCAAGGAAGGCTTAATTTGAAGATACTTTGAAATCTCGA	1505
Db	1445	TGATGAACAGAGAGACTTTCAAGGAAGGCTTAATTTGAAGATACTTTGAAATCTCGA	1504

QY	1506	CTCCAAAAATCAACGACAGCTTTAGCCGGATCATGACTATGTTGAATATGCAAGTTTGTG	1565
DB	1505	CTCCAAAAATCAACGACAGCTTTAGCCGGATCATGACTATGTTGAATATGCAAGTTTGTG	1564
QY	1566	TACGAGTGAGGAGATGGACAACCTCTGTGAAGAAATCTTCAAGGGTTATGGACCTCAAG	1625
DB	1565	TACGAGTGAGGAGATGGACAACCTCTGTGAAGAAATCTTCAAGGGTTATGGACCTCAAG	1624
QY	1626	GCCAAATGATCTACATTTGTTGAATCCAGTGCATCTTCTTGGGTACCCTGTGTGG	1685
DB	1625	GCCAAATGATCTACATTTGTTGAATCCAGTGCATCTTCTTGGGTACCCTGTGTGG	1684
QY	1686	ACAGATTAGAGATTTTACAGGACGAGGGCTCTACCTCTCAGACATCCCAATTCACATG	1745
DB	1685	ACAGATTAGAGATTTTACAGGACGAGGGCTCTACCTCTCAGACATCCCAATTCACATG	1744
QY	1746	CACGTAGGGATGTGGTCTTAATAGGGAAACAAGCCGAGCTCAAGATGGCGCTCAAGAAGA	1805
DB	1745	CACGTAGGGATGTGGTCTTAATAGGGAAACAAGCCGAGCTCAAGATGGCGCTCAAGAAGA	1804
QY	1806	GGCTGGGAAGCTGAAGGCTTACCTTGAGCAAGCCCAACCAAGCCCTGGAGGAGAGAAGA	1865
DB	1805	GGCTGGGAAGCTGAAGGCTTACCTTGAGCAAGCCCAACCAAGCCCTGGAGGAGAGAAGA	1864
QY	1866	AAAAAGACAGTAGACCTTCTGTGCTCCATATTTCCCTGTGAGGTGCTCAGCAGCTGTGGC	1925
DB	1865	AAAAAGACAGTAGACCTTCTGTGCTCCATATTTCCCTGTGAGGTGCTCAGCAGCTGTGGC	1924
QY	1926	AAGGGCAAGTTGTCAGCAAGCAAGAATTCAGTAAATGTCACCATGCTCTTCTCAGACATCG	1985
DB	1925	AAGGGCAAGTTGTCAGCAAGCAAGAATTCAGTAAATGTCACCATGCTCTTCTCAGACATCG	1984
QY	1986	TTGGGTTTCACTGCCATCTGCTCCAGAGCTCACCGCTGCAGGTGCTCACCATGCTCAATG	2045
DB	1985	TTGGGTTTCACTGCCATCTGCTCCAGAGCTCACCGCTGCAGGTGCTCACCATGCTCAATG	2044
QY	2046	CACGTGTACACTCGCTTCGACCAGCAGTGTGGAGAGCTGAGTGTCTCAAGGTTGGAGACCA	2105
DB	2045	CACGTGTACACTCGCTTCGACCAGCAGTGTGGAGAGCTGAGTGTCTCAAGGTTGGAGACCA	2104
QY	2106	TTGGGATGCTATTTGTGTAAGCTGGGGATACACAAAGAGATGATCATGCTGCTGTTTC	2165
DB	2105	TTGGGATGCTATTTGTGTAAGCTGGGGATACACAAAGAGATGATCATGCTGCTGTTTC	2164
QY	2166	AGATAGGGCTGATGGCCCTCAAGATGATGAGGTCTCTGATGAAGTTATGCTCCTCCATG	2225
DB	2165	AGATAGGGCTGATGGCCCTCAAGATGATGAGGTCTCTGATGAAGTTATGCTCCTCCATG	2224
QY	2226	GAGAACTATCAAGATCGGAATTTGGACTGCACCTTGATCGATCAGTTTTTGTGGGCTGTTG	2285
DB	2225	GAGAACTATCAAGATCGGAATTTGGACTGCACCTTGATCGATCAGTTTTTGTGGGCTGTTG	2284
QY	2286	GAGTTAAATGCCCGCTTACTGCTCTTTTGGAAACAATGTCCTCTGGCTTAACAAATTTTG	2345
DB	2285	GAGTTAAATGCCCGCTTACTGCTCTTTTGGAAACAATGTCCTCTGGCTTAACAAATTTTG	2344
QY	2346	AGTCTGCGAGTGTACCAAGAAAATCAATGTACGCCCAACAACCTTACAGATTTACTCAAAG	2405
DB	2345	AGTCTGCGAGTGTACCAAGAAAATCAATGTACGCCCAACAACCTTACAGATTTACTCAAAG	2404
QY	2406	ACTGCTCTGGTTCTGGTCTTTACCCCTCGATCAAGGGAGGAACCTTCCACCAAACTTCCCTA	2465
DB	2405	ACTGCTCTGGTTCTGGTCTTTACCCCTCGATCAAGGGAGGAACCTTCCACCAAACTTCCCTA	2464
QY	2466	GTGAAATCCCGGAATCTGCCATTTTCTGGATGCTTTACCAACAAGGAACAACTCAAAAC	2525
DB	2465	GTGAAATCCCGGAATCTGCCATTTTCTGGATGCTTTACCAACAAGGAACAACTCAAAAC	2524
QY	2526	CATGCTTCCAAAGAAGATGTGGAAGATGGCAATGCCATTTTTTATAGGCAAGCATCAG	2585
DB	2525	CATGCTTCCAAAGAAGATGTGGAAGATGGCAATGCCATTTTTTATAGGCAAGCATCAG	2584
QY	2586	GAATAGATTAGCAACCTATATACCTATTTTATAGTCTTTTGGGCTTTGACTCATTTGAAGAT	2645

QY 786 GGCTGAATGTTGCATCTTCAGAGAACATVTCGCAAGACACAAATAAAGAGAGAGGAAAT 845
DB 785 GGCTGAATGTTGCATCTTCAGAGAACATVTCGCAAGACACAAATAAAGAGAGGAAAT 844
QY 846 CTTTGGAAAGAGAGACTTTTGAAGAACAAATTCAGAGAACAGCAGTTCAGAGAGTTC 905
DB 845 CTTTGGAAAGAGAGACTTTTGAAGAACAAATTCAGAGAACAGCAGTTCAGAGAGTTC 904
QY 906 CAGTGGAGGTATCAAGAAATCTCTTGGTGAAGAGGTTTTTAAATATATTTACGAGGAAG 965
DB 905 CAGTGGAGGTATCAAGAAATCTCTTGGTGAAGAGGTTTTTAAATATATTTACGAGGAAG 964
QY 966 ATGAAACATCTCTTGGGGTGTGGAGGACCCCTTAAAGATTTTTTAAACAGCTTCAGTA 1025
DB 965 ATGAAACATCTCTTGGGGTGTGGAGGACCCCTTAAAGATTTTTTAAACAGCTTCAGTA 1024
QY 1026 CCCTTCGAAACAGAGCAGCATTCGCAAGACAGGAAAGGGGCGGCTTGAGGACG 1085
DB 1025 CCCTTCGAAACAGAGCAGCATTCGCAAGACAGGAAAGGGGCGGCTTGAGGACG 1084
QY 1086 CCTCCATTTCTATGCTGGATGAAGAGGATGATTTTCTACATGTTTACTACTTCTTCCCTA 1145
DB 1085 CCTCCATTTCTATGCTGGATGAAGAGGATGATTTTCTACATGTTTACTACTTCTTCCCTA 1144
QY 1146 AGAGAACCCCTCCCTGATTTCTCCGGCATATAAAGGAGCTGCTCAGTATATATG 1205
DB 1145 AGAGAACCCCTCCCTGATTTCTCCGGCATATAAAGGAGCTGCTCAGTATATATG 1204
QY 1206 AAAGGAAGTGGAGTCTGCTTAATGCTCCCTGCTTCCATATATGATGTCAGCGAGTTG 1265
DB 1205 AAAGGAAGTGGAGTCTGCTTAATGCTCCCTGCTTCCATATATGATGTCAGCGAGTTG 1264
QY 1266 TGAATCAGCCCTACTTGTGTTACTTCCGTTTACATGAAAGACCAAGCCATCCCTGTCCT 1325
DB 1265 TGAATCAGCCCTACTTGTGTTACTTCCGTTTACATGAAAGACCAAGCCATCCCTGTCCT 1324
QY 1326 CCAGCAAAACCCAGTCTCCTGCTGATTCCTCAGATGCTTCCATATGTCGCAAGACATTTCCAT 1385
DB 1325 CCAGCAAAACCCAGTCTCCTGCTGATTCCTCAGATGCTTCCATATGTCGCAAGACATTTCCAT 1384
QY 1386 TCATTTTCATGTTTGACAAAGATGATGACAAATTCGCAATTTGGCAATGGCATCAGAGGC 1445
DB 1385 TCATTTTCATGTTTGACAAAGATGATGACAAATTCGCAATTTGGCAATGGCATCAGAGGC 1444
QY 1446 TGATGAACAGAGAGACTTTTCAAGGAAGCCTTAATTTTGAAGAACTTTTGAATTTCTGA 1505
DB 1445 TGATGAACAGAGAGACTTTTCAAGGAAGCCTTAATTTTGAAGAACTTTTGAATTTCTGA 1504
QY 1506 CTCCAAAATCAACAGAGCTTTAGCGGATCATGACTATGTTGAATATCAGTTCGTTG 1565
DB 1505 CTCCAAAATCAACAGAGCTTTAGCGGATCATGACTATGTTGAATATCAGTTCGTTG 1564
QY 1566 TACAGTGAAGAGATGGGCAACTCTGTGAAGAACTTCAAGGGTTATGGACCTCAAG 1625
DB 1565 TACAGTGAAGAGATGGGCAACTCTGTGAAGAACTTCAAGGGTTATGGACCTCAAG 1624
QY 1626 GCCAAATGATCTACATTTGTAATCCAGTGCATCTTTGTTTTGGGGTACCCTGTGTGG 1685
DB 1625 GCCAAATGATCTACATTTGTAATCCAGTGCATCTTTGTTTTGGGGTACCCTGTGTGG 1684
QY 1686 ACAGATTAGAAGATTTTACAGGAGGAGGCTTACCTCTCAGACATCCCAATTCACANTG 1745
DB 1685 ACAGATTAGAAGATTTTACAGGAGGAGGCTTACCTCTCAGACATCCCAATTCACANTG 1744
QY 1746 CACTGAGGAGTGGTCTTAATAGGGAACAAGCCGAGCTCAAGATGGCTGAGAGGA 1805
DB 1745 CACTGAGGAGTGGTCTTAATAGGGAACAAGCCGAGCTCAAGATGGCTGAGAGGA 1804
QY 1806 GGCTGGGAAGCTGAAGGCTTACCTTGAAGCAAGCCACCAAGCCCTGGAGGAGGAAGA 1865
DB 1805 GGCTGGGAAGCTGAAGGCTTACCTTGAAGCAAGCCACCAAGCCCTGGAGGAGGAAGA 1864
QY 1866 AAAGACAGTAGACCTTCTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAGCTGTGGC 1925

DB 1865 AAAAGACAGTAGACCTTCTGTCTCCATATTTCCCTGTAGGTTGCTCAGCAGCTGTGC 1924
QY 1926 AAGGCAAGTTGTGCAAGCAAGTTCAGTATATGTCCACCATGCTCTTCTCAGACATCG 1985
DB 1925 AAGGCAAGTTGTGCAAGCAAGTTCAGTATATGTCCACCATGCTCTTCTCAGACATCG 1984
QY 1986 TTTGGGTTCACTGCATCTCTCCAGTGTCCAGGTCACCGCTGCAAGTTCATCACCATGCTCAATG 2045
DB 1985 TTTGGGTTCACTGCATCTCTCCAGTGTCCAGGTCACCGCTGCAAGTTCATCACCATGCTCAATG 2044
QY 2046 CACTGTACACTCTCGCTCGACAGTGTGGAGAGCTGAGAGCTGGATGTCTACAAGTGGAGACCA 2105
DB 2045 CACTGTACACTCTCGCTCGACAGTGTGGAGAGCTGAGAGCTGGATGTCTACAAGTGGAGACCA 2104
QY 2106 TTTGGCATGCTATTTGTGTAGCTGGGGATACACAAAGAGAGTGTACTCATGCTGTTC 2165
DB 2105 TTTGGCATGCTATTTGTGTAGCTGGGGATACACAAAGAGAGTGTACTCATGCTGTTC 2164
QY 2166 AGATAGCGCTGATGGCCCTGAAGATGATGAGCTCTCTGATGAAGTATGCTCCTCCCATG 2225
DB 2165 AGATAGCGCTGATGGCCCTGAAGATGATGAGCTCTCTGATGAAGTATGCTCCTCCCATG 2224
QY 2226 GAGAACCTATCAAGATGCGAATTTGGAGTGTGATGCTGTGATCAGTTTTTCTGGGCTGTG 2285
DB 2225 GAGAACCTATCAAGATGCGAATTTGGAGTGTGATGCTGTGATCAGTTTTTCTGGGCTGTG 2284
QY 2286 GAGTTAAATGCCCCGTACTGTCTTTTGGAAACAATGTCTCTGGCTTAACAAATTTG 2345
DB 2285 GAGTTAAATGCCCCGTACTGTCTTTTGGAAACAATGTCTCTGGCTTAACAAATTTG 2344
QY 2346 AGTCTGCAAGTGTACACGAAATAATCAATGTACAGCCCAACAATTTACAGATTTACTCAAAG 2405
DB 2345 AGTCTGCAAGTGTACACGAAATAATCAATGTACAGCCCAACAATTTACTCAAAG 2404
QY 2406 ACTGTCTGTTGTTGTTTACCCCTCGATCAAGGAGGAACTTCCACCAAACTTCCCTA 2465
DB 2405 ACTGTCTGTTGTTGTTTACCCCTCGATCAAGGAGGAACTTCCACCAAACTTCCCTA 2464
QY 2466 GTCAAAATCCCGGAATCTGCAATTTTCTGGATGCTTACCAACAAGGAACAATCTCAAAAC 2525
DB 2465 GTCAAAATCCCGGAATCTGCAATTTTCTGGATGCTTACCAACAAGGAACAATCTCAAAAC 2524
QY 2526 CATGCTTCCAAAAGAGATGTGGAAGATGGCAATGCCAATTTTTAGGCAAAAGCATCAG 2585
DB 2525 CATGCTTCCAAAAGAGATGTGGAAGATGGCAATGCCAATTTTTAGGCAAAAGCATCAG 2584
QY 2586 GAATAGATTAGCAACCTATATACCTATTTATAGTCTTTGGGGTTTGACTCATTTGAAGAT 2645
DB 2585 GAATAGATTAGCAACCTATATACCTATTTATAGTCTTTGGGGTTTGACTCATTTGAAGAT 2644
QY 2646 GTGTAGAGCTCTGAAAGCACTTTAGGGATTTAGATGGCTTAACAAGCAGTATTTAAATTT 2705
DB 2645 GTGTAGAGCTCTGAAAGCACTTTAGGGATTTAGATGGCTTAACAAGCAGTATTTAAATTT 2704
QY 2706 TCAGAGGCCAAGTCACAATCTTTCTCCTGTTTAAACATGACAAAATGTACTCTCAGTA 2765
DB 2705 TCAGAGGCCAAGTCACAATCTTTCTCCTGTTTAAACATGACAAAATGTACTCTCAGTA 2764
QY 2766 CTTTCAGCTCTTCAAGAAAAAACCCTTTAAAAGCTACTTTTGTGGGAGTATTTCT 2825
DB 2765 CTTTCAGCTCTTCAAGAAAAAACCCTTTAAAAGCTACTTTTGTGGGAGTATTTCT 2824
QY 2826 ATTATATAACAGCACTTACTACTGTACTTCAAAATTCAGCACCCTGTACATATATCAGA 2885
DB 2825 ATTATATAACAGCACTTACTACTGTACTTCAAAATTCAGCACCCTGTACATATATCAGA 2884
QY 2886 TAATTTAGTCAATTTTACAACTGATGGAGTCACTGCAATCTCATATCTCTGTTGGAAT 2945
DB 2885 TAATTTAGTCAATTTTACAACTGATGGAGTCACTGCAATCTCATATCTCTGTTGGAAT 2944
QY 2946 GCCATGTTTATTAAGTGTGTTTGTGATAGTGTCTGTCAAAAAATAAAAAAATAAAAAA 3005
DB 2945 GCCATGTTTATTAAGTGTGTTTGTGATAGTGTCTGTCAAAAAATAAAAAAATAAAAAA 3004

Db	2945	CCCATGGTTATTAAGTGTCTTTGTGNTAGTGTCTGCAAAAAAAGSTTTTGAATAGAA	3004
Qy	3006	AAAAA 3011	
Db	3005	ATTACA 3010	

RESULTS

RESULTS 3
ABV27568
ID ABV27568 standard; cDNA; 4742 BP.

DT 16-SEP-2002 (first entry)

Human prostate expression marker cDNA 27559.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW
pharmacogenomic marker; gene; ss.
KW

OS Homo sapiens.

PN WO200160860-A2.

23-AUG-2001.

20-FEB-2001: 2001WO-US051711

PR 17-FEB-2000: 2000US-183319P.

PK I6-MAR-2000; 2000US-189862P.
PB 25-MAY-2000: 2000US-207454P

PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.
XX

PA (MILL-) MILLENNIUM PREDICTIV
YY

PI Schlegel R, Endege WO, Mon

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule

PT for detecting presen

PS Claim 1; Page 5633-5634; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:

- (a) assessing whether a patient is afflicted with prostate cancer;
- (b) monitoring the progression of prostate cancer in a patient;
- (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
SQ Sequence 4742 BP: 1450 A: 933 C: 998 G: 1355 T: 6 other:
SQ

Qy 6 ATGCGATTGGGGGGCTGCAGAGACCAAGGACTCAGTTCCCTTCCCCTAGCTGTGAGCCTAG 65
 |||||
Dd 6 ATGCGCATTTGGGGGGCTGCAGAGACCAAGGACTCAGTTCCCTTCCCCTAGCTGTGAGCCTAG 65
 |||||

Db 1145 AGAAGACCACTCCCTCGATTCCTCCCGCATATAAAGCAGCTGCTACGTAATTATATG 1204
Qy 1206 AAACGGAAGTGGAAAGTGTGCTTAATGCTCCCTGCTTCCATATGATTCACGCGAGTTTG 1265
Db 1205 AAACGGAAGTGGAAAGTGTGCTTAATGCTCCCTGCTTCCATATGATTCACGCGAGTTTG 1264
Qy 1266 TGAATCAGCCCTACTTTGTTGTAATCCGTTTCAATGAAAGACCAAGCCATCCCTGTCOC 1325
Db 1265 TGAATCAGCCCTACTTTGTTGTAATCCGTTTCAATGAAAGACCAAGCCATCCCTGTCOC 1324
Qy 1326 CAGCAAAACCCAGTCCCTGCTGGTGCATTCACACATCGCTATTCGCAAGACATTTCCAT 1385
Db 1325 CAGCAAAACCCAGTCCCTGCTGGTGCATTCACACATCGCTATTCGCAAGACATTTCCAT 1384
Qy 1386 TCCATTTTCATGTTTGAAGAATATGACAAATCTGCAATTTGGCAATGGCATCAGAAGGC 1445
Db 1385 TCCATTTTCATGTTTGAAGAATATGACAAATCTGCAATTTGGCAATGGCATCAGAAGGC 1444
Qy 1446 TGATGAACAGGAGACTTTCAAGGAAAGCCTAAATTTTGAAGAATACATTTGAAATCTGA 1505
Db 1445 TGATGAACAGGAGACTTTCAAGGAAAGCCTAAATTTTGAAGAATACATTTGAAATCTGA 1504
Qy 1506 CTCCAAAATCAACACAGAGCTTTAGCGGATCATGACTATGTTGAATATGCAATTTGTTG 1565
Db 1505 CTCCAAAATCAACACAGAGCTTTAGCGGATCATGACTATGTTGAATATGCAATTTGTTG 1564
Qy 1566 TAGCAGTGAGGAGTGGGACAACTCTGTGAAGAAATCTTCAAGGTTATGGACCTCAAAAG 1625
Db 1565 TAGCAGTGAGGAGTGGGACAACTCTGTGAAGAAATCTTCAAGGTTATGGACCTCAAAAG 1624
Qy 1626 GCCAAATGATCTACATTTGTAATCCAGTGCAATCTGTTTTGGGGTCAACCTGTGTGG 1685
Db 1625 GCCAAATGATCTACATTTGTAATCCAGTGCAATCTGTTTTGGGGTCAACCTGTGTGG 1684
Qy 1686 ACAGATAGNAGATTTACAGACGAGGCTCTACCTCTCAGACATCCCAATTCAGATG 1745
Db 1685 ACAGATAGNAGATTTACAGACGAGGCTCTACCTCTCAGACATCCCAATTCAGATG 1744
Qy 1746 CACTGAGGATGTGTCTTAATAGGGAACAAAGCCGAGCTCAAGATGGCTTGAAGAAGA 1805
Db 1745 CACTGAGGATGTGTCTTAATAGGGAACAAAGCCGAGCTCAAGATGGCTTGAAGAAGA 1804
Qy 1806 GCGTGGGGAAGCTGAAGGCTACCTTTAGCAAGACCCCAAGCCCTGGAGGAGGAGAAGA 1865
Db 1805 GCGTGGGGAAGCTGAAGGCTACCTTTAGCAAGACCCCAAGCCCTGGAGGAGGAGAAGA 1864
Qy 1866 AAAAGACAGTAGACCTTCTGTCTCCATATTTCCCTGTGAGCTGCTCAGACGCTGGC 1925
Db 1865 AAAAGACAGTAGACCTTCTGTCTCCATATTTCCCTGTGAGCTGCTCAGACGCTGGC 1924
Qy 1926 AAGGCAAGTTGTGCAAGCAAGAAGTTCAAGTAATGTCCACCATGCTTCTCAGACATCG 1985
Db 1925 AAGGCAAGTTGTGCAAGCAAGAAGTTCAAGTAATGTCCACCATGCTTCTCAGACATCG 1984
Qy 1986 TTGGGTTCACTGCCATCTGCTCCAGTGTCCAGGTCACCGTGCAGGTATCACCATGCTCAATG 2045
Db 1985 TTGGGTTCACTGCCATCTGCTCCAGTGTCCAGGTCACCGTGCAGGTATCACCATGCTCAATG 2044
Qy 2046 CACTGTACACTCGCTTCGACACGACGCTGGAGAGCTGGATGCTACAGAGTGGAGACCA 2105
Db 2045 CACTGTACACTCGCTTCGACACGACGCTGGAGAGCTGGATGCTACAGAGTGGAGACCA 2104
Qy 2106 TTGGCGATGCTATTGTAGCTGGGGGATTACAAAAGAGAGTGTACTCATGTCTGTTTC 2165
Db 2105 TTGGCGATGCTATTGTAGCTGGGGGATTACAAAAGAGAGTGTACTCATGTCTGTTTC 2164
Qy 2166 AGATAGCGGTGATGGCCCTGAAGATGATGAGCTCTCTGATGAAGTATGATGCTCCCATG 2225
Db 2165 AGATAGCGGTGATGGCCCTGAAGATGATGAGCTCTCTGATGAAGTATGATGCTCCCATG 2224
Qy 2226 GAGAACTATCAAGATGCGAATTTGGACTGCACTCTGGATCAGTTTGTGCGCGCTGTTG 2285

Db 2225 GAGAACTATCAAGATGCGAATTTGGACTGCACTCTGGATCAGTTTGTGCTGGCGCTGTTG 2284
Qy 2286 GAGTAAATGCCCGCTTACTGTCTTTTGGAAACAATGTCACTCTGGCTTAACAAAATTG 2345
Db 2285 GAGTAAATGCCCGCTTACTGTCTTTTGGAAACAATGTCACTCTGGCTTACAAAATTG 2344
Qy 2346 AGTCTCTCAGTGTACCAGGAAATCAATGTACGCCCAACAACTTACAGATTTACTCAAG 2405
Db 2345 AGTCTCTCAGTGTACCAGGAAATCAATGTACGCCCAACAACTTACAGATTTACTCAAG 2404
Qy 2406 ACTGCTCTGTTTCTGTTTACCCCTCGATCAAGGGAGGAACTTCCACCAATTCCTTA 2465
Db 2405 ACTGCTCTGTTTCTGTTTACCCCTCGATCAAGGGAGGAACTTCCACCAATTCCTTA 2464
Qy 2466 GTGAAATCCCCGGAATCTGCCATTTCTGGATGCTTACCAACAGAGAACAACTCAAAAC 2525
Db 2465 GTGAAATCCCCGGAATCTGCCATTTCTGGATGCTTACCAACAGAGAACAACTCAAAAC 2524
Qy 2526 CATGCTTCCAAAAGAAAGATGTGAAGATGGCAATGCCAATTTTTAGGCAAAAGCATCAG 2585
Db 2525 CATGCTTCCAAAAGAAAGATGTGAAGATGGCAATGCCAATTTTTAGGCAAAAGCATCAG 2584
Qy 2586 GAATAGATTAGCAACCTATATACCTATTTAAGTCTTTGGGGTTTGACTCATTTGAAGAT 2645
Db 2585 GAATAGATTAGCAACCTATATACCTATTTAAGTCTTTGGGGTTTGACTCATTTGAAGAT 2644
Qy 2646 GTGTAGAGCTCTGAAAGCACTTTAGGATTTAGATGGCTTAACAAGCAGTATTAAATTT 2705
Db 2645 GTGTAGAGCTCTGAAAGCACTTTAGGATTTAGATGGCTTAACAAGCAGTATTAAATTT 2704
Qy 2706 TCAGAGCAAGTCACAATCTTCTCTGTTTAACTGACAAAATGTACTACTTCAAGTA 2765
Db 2705 TCAGAGCAAGTCACAATCTTCTCTGTTTAACTGACAAAATGTACTACTTCAAGTA 2764
Qy 2766 CTTGAGCTCTTCAAGAAAAAACCCTTAAAAAGCTACTTTTGTGGGAGTATTCT 2825
Db 2765 CTTGAGCTCTTCAAGAAAAAACCCTTAAAAAGCTACTTTTGTGGGAGTATTCT 2824
Qy 2826 ATTATATAACAGCACTTACTTACCTGTACTCATAAATTCAGCACCCTGTACATATATCAGA 2885
Db 2825 ATTATATAACAGCACTTACTTACCTGTACTCATAAATTCAGCACCCTGTACATATATCAGA 2884
Qy 2886 TAATTTAGTCAATTTGACAACTGTGGAGTCACTTGCATCTCATATCTCTGTTGGAAT 2945
Db 2885 TAATTTAGTCAATTTGACAACTGTGGAGTCACTTGCATCTCATATCTCTGTTGGAAT 2944
Qy 2946 GCCATGTTTATTAAGTGTGTTTGTGATAGTGTCTCAAAAAAAGGTTTGTGANTAGAA 3005
Db 2945 GCCATGTTTATTAAGTGTGTTTGTGATAGTGTCTCAAAAAAAGGTTTGTGANTAGAA 3004
Qy 3006 AAAAAA 3011
Db 3005 ATTACA 3010

RESULT 6
ABK92193
ID ABK92193 standard; DNA; 3004 BP.
XX AC ABK92193;
XX XX
DF 15-AUG-2002 (first entry)
XX Prostate cancer-associated DNA sequence #79.
DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW gene therapy; gene; ds.
XX Mammalia.
XX WO200230268-A2.
XX PN
XX 18-APR-2002.
PD

1374	Db	TCCATTCCATTTCATGTTTGACAAGATATGCAAAATTCGTGCAATTTGGCAATTTGGCAATCGC	1433
1441	QY	AAGGCTGATGAACAGAGAGACATTTTCAAGGAAAGCCTAATTTTGAAGAAATACTTTTGA	1500
1434	Db	AAGGCTGATGAACAGAGAGACTTTCAGGAAGCCTAATTTT--GAATACTTTTGAAT	1490
1501	QY	TTGTACTCCAAAATCAACCAGACGTTTATGCGGATCATGTACTATGTTGAATATGCA	1560
1491	Db	TTGTACTCCAAAATCAACCAGACCTTTAGCGGATCATGTACTATGTTGAATATGCA	1550
1561	QY	TGTTGTACGAGTCAGAGATGGGACAACCTCTGTGAAGAAATCTTCAAGGGTTATGGACCT	1620
1551	Db	TGTTGTACGAGTCAGAGATGGGACAACCTCTGTGAAGAAATCTTCAAGGGTTATGGACCT	1610
1621	QY	CAAAGGCCAAATGATCTACATCTGTTCAATCCAGTGCAAATCTTGTCTTTTGGGGTCA	1680
1611	Db	CAAAGGCCAAATGATCTACATCTGTTCAATCCAGTGCAAATCTTGTCTTTTGGGGTCA	1670
1681	QY	TGTGGACAGATTAGAAGATTTTACAGGACGAGGGCTCTACCTCTTCAGACATCCC	1740
1671	Db	TGTGGACAGATTAGAAGATTTTACAGGACGAGGGCTCTACCTCTTCAGACATCCC	1730
1741	QY	CAATGCACGTAGGGATGTGGTCTTAAATAGGGGACAAGCCCGAGCTCAAGATGG	1800
1731	Db	CAATGCACGTAGGGATGTGGTCTTAAATAGGGGACAAGCCCGAGCTCAAGATGG	1790
1801	QY	GAAGAGGCTGGGGAAGCTGAAGGCTACCTCTGAGCAAGCCCAAGAGCCCTGGAGG	1860
1791	Db	GAAGAGGCTGGGGAAGCTGAAGGCTACCTCTGAGCAAGCCCAAGAGCCCTGGAGG	1850
1861	QY	GAAGAAAAACAGTAGTAGACCTTCTGTGCTCCATATTTCCCTGTGAGGTTGCTCAG	1920
1851	Db	GAAGAAAAACAGTAGTAGACCTTCTGTGCTCCATATTTCCCTGTGAGGTTGCTCAG	1910
1921	QY	GTGGCAAGGCCAAGTTGTGCAAGCCCAAGAAAGTTTCAGTAATGTCACCATGCTCT	1980
1911	Db	GTGGCAAGGCCAAGTTGTGCAAGCCCAAGAAAGTTTCAGTAATGTCACCATGCTCT	1970
1981	QY	CATCGTTGGGTTTCACATGCCATCTGCTCCAGTGCTCACCGCTGCAGGTCATCAC	2040
1971	Db	CATCGTTGGGTTTCACATGCCATCTGCTCCAGTGCTCACCGCTGCAGGTCATCAC	2030
2041	QY	CAATGCACGTACATCTGGTTTGCAACGACAGTGTGGAGAGCTGGATGCTTACAAG	2100
2031	Db	CAATGCACGTACATCTGGTTTGCAACGACAGTGTGGAGAGCTGGATGCTTACAAG	2090
2101	QY	GACCAATGGCGATGCCCTATTGTGTAGC--TGGGGGATTCACAAAGAGAGTGATCT	2159
2091	Db	GACCAAT--GCGATGCGCTATTGTGTGGCTTTGGGGGATTCACAAAGAGAGTGAT	2149
2160	QY	CTGTTCCAGATAGGGCTGATGGCCCTCAAGATGATGGAGCTCTCTGATGAAGTTAT	2219
2150	Db	CTGTTCCAGATAGGGCTGATGGCCCTCAAGATGATGGAGCTCTCTGATGAAGTTAT	2209
2220	QY	CCCATGGGAACCTATCAAGATGCGAATTTGGAGCTGCACCTCTGGATCAGTTTTG	2279
2210	Db	CCCATGGGAACCTATCAAGATGCGAATTTGGAGCTGCACCTCTGGATCAGTTTTG	2269
2280	QY	TGTTTGGAGTTAAATGCCCCGGTTACTGTCTTTTGGAAACAAATGTCTACTCTG	2339
2270	Db	TGTTTGGAGTTAAATGCCCCGGTTACTGTCTTTTGGAAACAAATGTCTACTCTG	2329
2340	QY	AATTTTCAGTCTGCAAGTGTACACAGAAAAATCAATGTCAGGCCCAACAACCTT	2399
2330	Db	AATTTTCAGTCTGCAAGTGTACACAGAAAAATCAATGTCAGGCCCAACAACCTT	2389
2400	QY	TCAAGAGACTGTCTTGGTTTTCGTTTACCCCTCGATCAAGGGAGGAACCTTCC	2459
2390	Db	TCAAGAGACTGTCTTGGTTTTCGTTTACCCCTCGATCAAGGGAGGAACCTTCC	2449
2460	QY	TCCCTAGTGAATTCGCCGGAAATCTGCCATTTTCTGGATGCTTACCAACAGGA	2519
2450	Db	TCCCTAGTGAATTCGCCGGAAATCTGCCATTTTCTGGATGCTTACCAACAGGA	2509

XX WPI: 2002-471335/50.
DR P-PSDB; ABG61940.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes
PT are expressed in a prostate tissue
XX
PS Claim 22; Page 421; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate
CC cancer-associated transcript in a cell from a patient. The method
CC comprises contacting a biological sample from the patient with
CC prostate cancer-associated polynucleotides (designated PC genes) that
CC selectively hybridise to a sequence that is at least 80% identical
CC to them. The prostate cancer-associated polynucleotide sequences
CC are differentially expressed in prostate tumour tissue or in
CC prostate cancer and are derived from the tissues of various
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC The methods of the invention are useful for diagnosing and treating
CC prostate cancer in mammals. The prostate cancer-associated genes are
CC useful for diagnosing or treating prostate cancer, as well as for
CC identifying modulators of prostate cancer or agents that inhibit
CC prostate cancer. The nucleic acid sequences are particularly useful
CC in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences.
XX
SQ Sequence 3004 BP; 874 A; 679 C; 710 G; 741 T; 0 other;
Query Match 94.7%; Score 2855.4; DB 24; Length 3004;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2976; Conservative 0; Mismatches 26; Indels 15; Gaps 9;
QY 1 CCCTATGGGATTTGGCGCTGCAGACACGAGCTCAGTCCCTGCCCTAGTCTGAG 60
DB 1 CCCTATGGGATTTGGCGCTGCAGACACGAGCTCAGTCCCTGCCCTAGTCTGAG 60
QY 61 CTTAGTGGGTTGGGACTCAGCTCAGAGTCTAGTCTTCCAGAGAGGTTTCAGTCCAGAGTT 120
DB 61 CTTAGTGGGTTGGGACTCAGCTCAGAGTCTAGTCTTCCAGAGAGGTTTCAGTCCAGAGTT 120
QY 121 TTCTACACTTTTCTTCCGCTAGACGAGCAGGAGCTTGAACAGACCCAGCGGAGGAC 180
DB 121 TTCTACACTTTTCTTCCGCTAGACGAGCAGGAGCTTGAACAGACCCAGCGGAGGAC 180
QY 181 ACCTGTGGGGAGGGAGCGCTTGGAGAGCTTAGAGACCCAGCGGCGGTGATCTCACC 240
DB 181 ACCTGTGGGGAGGGAGCGCTTGGAGAGCTTAGAGACCCAGCGGCGGTGATCTCACC 240
QY 241 ATGTGGGGATTTGGAGGGCGCCCTGGAGCTGCTAGAGATCCGGAAGCACAGCCCGGAG 300
DB 241 ATGTGGGGATTTGGAGGGCGCCCTGGAGCTGCTAGAGATCCGGAAGCACAGCCCGGAG 300
QY 301 GTCTGCGAAGCCACAGACTCGGCTCTTGGAGAAAGCTGAGCAGGGGGCCACCGCGG 360
DB 301 GTCTGCGAAGCCACAGACTCGGCTCTTGGAGAAAGCTGAGCAGGGGGCCACCGCGG 360
QY 361 TCTCCGCGCTGTCTGCACCTGTGCGCTGAGCTGAGTGCATGACATGACATCCAG 420
DB 361 TCTCCGCGCTGTCTGCACCTGTGCGCTGAGCTGAGTGCATGACATGACATCCAG 420
QY 421 TTACCAAGTGTCTTGAATGATAGTGGCTTCTGTTTGTGAGTCTCATATAAGAACTACAG 480
DB 421 TTACCAAGTGTCTTGAATGATAGTGGCTTCTGTTTGTGAGTCTCATATAAGAACTACAG 480
QY 481 CTATCAGGAGGAGATCCGAGAGGTTAGAGACACCAACCATGTTCTGCACAGAGCT 540
DB 481 CTATCAGGAGGAGATCCGAGAGGTTAGAGACACCAACCATGTTCTGCACAGAGCT 540
QY 541 CAAGGATCTCAAGATCACAGGAGTGTCTCTTCTACTGCGCAGGCTCAAGTTCC 600
DB 541 CAAGGATCTCAAGATCACAGGAGTGTCTCTTCTACTGCGCAGGCTCAAGTTCC 600

DB 540 CAAGGATCTCAAGATCACAGGAGAGTGTCTCTTCTACTGGCACCAGGTCAAGTTCC 599
QY 601 TAACAGAGTCTTTCAGAGGAGGAGCAGGAGAGCTCAGAGAGCTGCAAAAGCAACCGTGCCTCAT 660
DB 600 TAACAGAGTCTTTCAGAGGAGGAGCAGGAGAGCTCAGAGAGCTGCAAAAGCAACCGTGCCTCAT 659
QY 661 CTGCTAAGACATTTCTTCTGAGAGAACATACAAAGAGTCTTCTTCAAGAAAAACAGTCTG 720
DB 660 CTGCTAAGACATTTCTTCTGAGAGAACATACAAAGAGTCTTCTTCAAGAAAAACAGTCTG 719
QY 721 GAGCCGAGTCTATCTTTCACACTTTGGCAGAGAGTATTGTCAAACTGATTTTCCACAGTT 780
DB 720 GAGCCGAGTCTATCTTTCACACTTTGGCAGAGAGTATTGTCAAACTGATTTTCCACAGTT 779
QY 781 TGAACGGCTGAATTTGTCACACTTTCAGAGAACATTTGGCAAGCACAATAAAGAAAGCAG 840
DB 780 TGAACGGCTGAATTTGTCACACTTTCAGAGAACATTTGGCAAGCACAATAAAGAAAGCAG 839
QY 841 GAATCTTTGGAAGAGAGACACTTTGAAAAACAATTCGAGAGCAAGCAGTTGCGAGGAG 900
DB 840 GAATCTTTGGAAGAGAGACACTTTGAAAAACAATTCGAGAGCAAGCAGT---GCAGC 895
QY 901 AGTTCCAGTGGAGTTTATCAAGAAATCTCTTGTGAAGAGTCTTTTAAATATGTTACGA 960
DB 896 AGAGTCCAGTGGAGTTATCAAGAAATCTCTTGTGAAGAGTCTTTTAAATATGTTACGA 955
QY 961 GGAAGATGAACACATCTTGGGGTGGTGGAGCACCCTTAAAGATTTTAAACAGCTT 1020
DB 956 GGAAGATGAACACATCTTGGGGTGGTGGAGCACCCTTAAAGA-TTTTAAACAGCTT 1014
QY 1021 CAGTACCTTCTCAACAGAGCAGCCATTTCCCAAGAGCAGGAAAAAGGGCAGGCTTGA 1080
DB 1015 CAGTACCTTCTTCAACAGAGCAGCCATTTCCCAAGAGCAGGAAAAAGGGCAGGCTTGA 1073
QY 1081 GGACGGCTCCATCTATGCTGATGAAGAGAGTATTTTCTACATGTTTACTACTTCTT 1140
DB 1074 GGACGGCTCCATCTATGCTGATGAAGAGAGTATTTTCTACATGTTTACTACTTCTT 1133
QY 1141 CCCTAAGAGAACACCTCCCTGATCTTCCCGGCATCATAAAGCAGCTGCTCAGCTATT 1200
DB 1134 CCCTAAGAGAACACCTCCCTGATCTTCCCGGCATCATAAAGCAGCTGCTCAGCTATT 1193
QY 1201 ATATGAACGGAAGTGGAGTGCCTGATTAATGCTTCCCTGCTTCCATATGATTTGACGCGA 1260
DB 1194 ATATGAACGGAAGTGGAGTGCCTGATTAATGCTTCCCTGCTTCCATATGATTTGACGCGA 1253
QY 1261 GTTTGTGAATCAGCCCTACTTGTGTACTCCGTTCACATGAAAGCAGCAGGAGCAGCT 1320
DB 1254 GTTTGTGAATCAGCCCTACTTGTGTACTCCGTTCACATGAAAGCAGCAGGAGCAGCT 1313
QY 1321 GTCCCGCAGAAACCCAGTCTCGCTGCTGATTTCCCATCATGCTATTTCTGCAAGACATT 1380
DB 1314 GTCCCGCAGAAACCCAGTCTCGCTGCTGATTTCCCATCATGCTATTTCTGCAAGACATT 1373
QY 1381 TCCATTTCCATTTTCTGCAAGAGATGACAAATTTCTGCAATTTGGCAATGGCATCAG 1440
DB 1374 TCCATTTCCATTTTCTGCAAGAGATGACAAATTTCTGCAATTTGGCAATGGCATCAG 1433
QY 1441 AAGGCTGATGAACAGAGAGACTTTTCAAGGAAAGCCTTAATTTTGAAGAATACTTTGAAT 1500
DB 1434 AAGGCTGATGAACAGAGAGACTTTTCAAGGAAAGCCTTAATTTT---GAATACTTTGAAT 1490
QY 1501 TCTGACTTCCAAAATCAACAGAGCTTTAGCGGGATCATGACTATGTTGAATATGCAAGTT 1560
DB 1491 TCTGACTTCCAAAATCAACAGAGCTTTTAGCGGGATCATGACTATGTTGAATATGCAAGTT 1550
QY 1561 TGTGTACGAGTGAAGAGATGGGACAACTCTGTGAAGAAATCTTCAAGGGTATGAGACCT 1620
DB 1551 TGTGTACGAGTGAAGAGATGGGACAACTCTGTGAAGAAATCTTCAAGGGTATGAGACCT 1610
QY 1621 CAAAGGCCAAATGATCTACATTTGTAATCCAGTGGCAATCTTGTGTTTGGGTCACCTG 1680
DB 1611 CAAAGGCCAAATGATCTACATTTGTAATCCAGTGGCAATCTTGTGTTTGGGTCACCTG 1670

QY	1681	TGTGGACAGATTAGAACATTTTACAGGACGAGGGCTCTACTCTCAGACATCCCAATTCA	1740
DB	1671	TGTGGACAGATTAGAACATTTTACAGGACGAGGGCTCTACTCTCAGACATCCCAATTCA	1730
QY	1741	CAATGCACCTCAGGGATGTGCTCTTAATAGGGGAACAAGCCGAGCTCAAGATGSCCTGAA	1800
DB	1731	CAATGCACCTCAGGGATGTGCTCTTAATAGGGGAACAAGCCGAGCTCAAGATGSCCTGAA	1790
QY	1801	GAAGAGGCTGGGAAGCTGAAGGCTACCCCTTGAGCAAGCCACCAAGCCCTGGAGGAGGA	1860
DB	1791	GAAGAGGCTGGGAAGCTGAAGGCTACCCCTTGAGCAAGCCACCAAGCCCTGGAGGAGGA	1850
QY	1861	GAAGAAAAACACAGTAGACACCTCTGTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAGCT	1920
DB	1851	GAAGAAAAACACAGTAGACACCTCTGTGCTCCATTTTCCCTGTGAGGTTGCTCAGCAGCT	1910
QY	1921	GTGGCAAGGCAAGTTGTGTCAAGCCAAAGAAGTTTCAGTAAATGTCACCATGCTCTTCTCAGA	1980
DB	1911	GTGGCAAGGCAAGTTGTGTCAAGCCAAAGAAGTTTCAGTAAATGTCACCATGCTCTTCTCAGA	1970
QY	1981	CATCGTTGGGTTCCACTGCCATCTGCTCCCACTGCTCACCGCTCAGGTCATCACCATGCT	2040
DB	1971	CATCGTTGGGTTCCACTGCCATCTGCTCCCACTGCTCACCGCTCAGGTCATCACCATGCT	2030
QY	2041	CAATGCACTGTACACTCGCTTTCACCAAGCAGTGTGGAGAGCTGGATGCTTACAAAGTGGA	2100
DB	2031	CAATGCACTGTACACTCGCTTTCACCAAGCAGTGTGGAGAGCTGGATGCTTACAAAGTGGA	2090
QY	2101	GACCAATTGGCGATGCCATTATGTGTAGC - TGGGGGATTACAAAGACAGTGAATCTCATG	2159
DB	2091	GACCAAT - GCGATGCGCTATTTGTGTGGCTTGGGGGATTACAAAGACAGTGAATCTCATG	2149
QY	2160	CTGTTCCAGATAGCGCTGATGGCCCTGAAGATGATGGAGCTCTCTGATGAAGTTATGTCCTC	2219
DB	2150	CTGTTCCAGATAGCGCTGATGGCCCTGAAGATGATGGAGCTCTCTGATGAAGTTATGTCCTC	2209
QY	2220	CCCATGGAGAACCTATCAAGATCGGATTCGACTGCACCTCGGATCAGTTTTCCTGGCGG	2279
DB	2210	CCCATGGAGAACCTATCAAGATCGGATTCGACTGCACCTCGGATCAGTTTTCCTGGCGG	2269
QY	2280	TCGTTGGAGTTAAATGCCCGCTTACTGTCTTTTTGGAAACAATGTCACTCTGGCTAACCA	2339
DB	2270	TCGTTGGAGTTAAATGCCCGCTTACTGTCTTTTTGGAAACAATGTCACTCTGGCTAACCA	2329
QY	2340	AAATTTGAGTCTGCAGTGTACACAGAAAAATCAATGTCAGCCCAACAACCTACAGATTAC	2399
DB	2330	AAATTTGAGTCTGCAGTGTACACAGAAAAATCAATGTCAGCCCAACAACCTACAGATTAC	2389
QY	2400	TCAAAGACTCTCTGGTTTCGTGTTTACCCCTCGATCAAGGGAGGAACTTCCACCAAACT	2459
DB	2390	TCAAAGACTCTCTGGTTTCGTGTTTACCCCTCGATCAAGGGAGGAACTTCCACCAAACT	2449
QY	2460	TCCTTAGTGAATTCGCCGGAATTCGCCATTTTCTGGATGCTTACCACCAAGGAACAACT	2519
DB	2450	TCCTTAGTGAATTCGCCGGAATTCGCCATTTTCTGGATGCTTACCACCAAGGAACAACT	2509
QY	2520	CAAAACCATGCTTCCAAAGAAAGATGTGGAAGATGGCAATGCGCAATTTTATAGGCAAG	2579
DB	2510	CAAAACCATGCTTCCAAAGAAAGATGTGGAAGATG - -CAAGGCAATTTTATAGGCAAG	2567
QY	2580	CATCAGAAATAGATTAGCAACCTATATACCTATTTATAGCTCTTGGGGTTTCACATCAT	2639
DB	2568	CATCAGAAATAGATTAGCAACCTATATACCTATTTATAGCTCTTGGGGTTTCACATCAT	2627
QY	2640	GAAGATGTGTAGAGCCCTCTGAAAGCACTTTAGGGATTTAGATGGCTTAAACAAGCATATT	2699
DB	2628	GAAGATGTGTAGAGCCCTCTGAAAGCACTTTAGGGATTTAGATGGCTTAAACAAGCATATT	2687
QY	2700	AAAATTTCCAGAGCCCAAGTCACAATCTTCTCCTGTTTAAACATGACAAAAATGACTCACT	2759
DB	2688	AAAATTTCCAGAGCCCAAGTCACAATCTTCTCCTGTTTAAACATGACAAAAATGACTCACT	2747

RESULT 8

AAZ51684

AAZ51004
ID AAZ51684 standard: CDNA: 2715 BP.

AA
AC AA251684:XX
DT 04-JUL-2000 (first entry)XX
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708

DE Human cyclic nucleotide-associated protein-3 (CNAP-3) cDNA.

Cyclic nucleotide-associated protein-3; CNAP-3; human; cytostatic;
anti-arteriosclerotic; hepatotropic; anti-leukemic; anti-inflammatory;
immunomodulatory; anti-asthmatic; anti-anemic; anti-diabetic; diagnosis;
anti-sclerotic; dermatological; neuroprotective; anti-epileptic; cancer;
anti-Alzheimer's; anti-Parkinsonian; cerebroprotective; ophthalmological;
anti-infertility; anti-allergic; vasotropic; immunosuppressive;
hypotensive; gene therapy; prevention; treatment; arteriosclerosis;
cell proliferative disorder; autoimmune/inflammatory; diabetes mellitus;
neurological; vision; reproductive; smooth muscle; ss.

OS Homo sapiens.

XX	Key	Location/Qualifiers
FH		

in key	location/quad/rails
FT	440.2512
CDS	

```
FT      /*tag= a
```

FT /product= "Human CNAP-3 protein"

FT Shares 89% identity to human soluble guanylate

FT cyclase large subunit"

misc_binding

```

F1T b
bound
/*tag=

```

```

F01 /bound_molety= "Primer or probe
F01 /note="Useful for amplification
F01

```

FT **FT**

/note- user! for amplification of hybridisation techniques"

17 XX

PN WO200014248-A1.

PD 16-MAR-2000.

XX XX

PF 03-SEP-1999; 99WO-US20287.

XX

PR 04-SEP-1998; 98US-0148904.

XX
TIVW-1 TIVVED TIVVDM TIVV
TIVV

PA (INCY-) INCYTE PHARM INC.

XX Hillman, JI. V110 H C11001

PI Hillman JL, Yue H, Guegler NJ, Corley NC, Patterson C, Tang YI, XX

WPI: 2000-256994/22.

DR P-PSDB; AAY70475.

XX

Isolated cyclic nucleotide associated proteins useful for preventing, diagnosing and treating cell proliferative, autoimmune/inflammatory, neurological, vision, reproductive and smooth muscle disorders -

XX	Example 3; Page 71-72; 78pp; English.	
PS	The present sequence is the cDNA encoding human cyclic nucleotide	
XX	associated protein-3 (CNAP-3), identified in Incyte clone 159278,	
CC	that is isolated from ADENINB01 cDNA library. It is expressed in	
CC	nervous, reproductive, cardiovascular and developmental tissues.	
CC	CNAP sequences may be used for prevention, treatment and diagnosis of	
CC	diseases associated with altered CNAP expression such as, cell	
CC	proliferative disorders (e.g. arteriosclerosis, cirrhosis, leukemia,	
CC	lymphoma and cancer of the breast, prostate, lung and brain), autoimmune/	
CC	inflammatory disorders (e.g. asthma, anaemia, diabetes mellitus, multiple	
CC	sclerosis and psoriasis), neurological disorders (e.g. epilepsy,	
CC	Alzheimer's/Parkinson's disease and strokes), vision disorders (e.g.	
CC	conjunctivitis, glaucoma, cataracts and retinitis pigmentosa),	
CC	reproductive disorders (e.g. infertility, uterine fibroids, ectopic	
CC	pregnancies and impotence) and smooth muscle disorders (e.g. angina,	
CC	anaphylactic shock, Kearns-Sayre syndrome and hypertension). It can also	
CC	be used for gene therapy.	
XX		
SQ	Sequence 2715 BP; 799 A; 604 C; 662 G; 650 T; 0 other;	
	Query Match 78.5%; Score 2368; DB 21; Length 2715;	
	Best Local Similarity 99.8%; Pred. No. 0;	
	Matches 2382; Conservative 0; Mismatches 0; Indels 4; Gaps 1;	
QY	412 ACATCCACGTTACCAAGTGTCTTGAATGATAGTGGCTTCTGTTGTCTAGTCTCATATAA 471	
DB		
QY	328 ACATCCACGTTACCAAGTGTCTTGAATGATAGTGGCTTCTGTTGTCTAGTCTCATATAA 387	
DB		
QY	472 GAACCTACGCTCATCAGGAGGATCGCAGCAGGGAAGAGACACCAACACATGTTCTG 531	
DB		
QY	388 GAACCTACGCTCATCAGGAGGATCGCAGCAGGGAAGAGACACCAACACATGTTCTG 447	
DB		
QY	532 CACGAAGCTCAAGGATCTCAAGATCACAGGAGTGTCTTCTCTTACTGSCACCG 591	
DB		
QY	448 CACGAAGCTCAAGGATCTCAAGATCACAGGAGTGTCTTCTCTTACTGSCACCG 507	
DB		
QY	592 TCAAGTTCCTAACAGGATCTTCAAGGAGGAGCAGGAGGATCTTCTCTTACTGSCACCG 651	
DB		
QY	508 TCAAGTTCCTAACAGGATCTTCAAGGAGGAGCAGGAGGATCTTCTCTTACTGSCACCG 567	
DB		
QY	652 CGTGCCCATCTCAGAGATCTTCAAGGAGGAGCAGGAGGATCTTCTCTTACTGSCACCG 711	
DB		
QY	568 CGTGCCCATCTCAGAGATCTTCAAGGAGGAGCAGGAGGATCTTCTCTTACTGSCACCG 627	
DB		
QY	712 AACAGTTCGAGCGGAGTCTATCTTCAAGGAGGAGCAGGAGGATCTTCTCTTACTGSCACCG 771	
DB		
QY	628 AACAGTTCGAGCGGAGTCTATCTTCAAGGAGGAGCAGGAGGATCTTCTCTTACTGSCACCG 687	
DB		
QY	772 CCCAGAGTTGAACGGCTGAATGTTGCACTTTCAGAGAACATTTGGCAAGACCAAAATAA 831	
DB		
QY	688 CCCAGAGTTGAACGGCTGAATGTTGCACTTTCAGAGAACATTTGGCAAGACCAAAATAA 747	
DB		
QY	832 AGAAGCAGGAATCTTTCGAAGAGAGACATTTGAAAACCAATTCAGAGCAAGCAGT 891	
DB		
QY	748 AGAAGCAGGAATCTTTCGAAGAGAGACATTTGAAAACCAATTCAGAGCAAGCAGT 807	
DB		
QY	892 TGCAGCAGGAGTTCAGTGGAGGTATCAAGAAATCTTGTGTAAGAGGTTTTTAAAT 951	
DB		
QY	808 TGCAGCAGGAGTTCAGTGGAGGTATCAAGAAATCTTGTGTAAGAGGTTTTTAAAT 867	
DB		
QY	952 ATGTTACAGGAGGATGAAGACATCTTGGGGTGGTGGAGGACCCCTTAAAGATTTTTT 1011	
DB		
QY	868 ATGTTACAGGAGGATGAAGACATCTTGGGGTGGTGGAGGACCCCTTAAAGATTTTTT 927	
DB		
QY	1012 AAACAGCTTCAGTACCCTTCTGAACAGAGCAGCCATTCGCCAAGAGCAGGAAAAGGGG 1071	
DB		
QY	928 AAACAGCTTCAGTACCCTTCTGAACAGAGCAGCCATTCGCCAAGAGCAGGAAAAGGGG 987	
DB		
QY	1072 CAGGCTTGAGGAGCGCTCCATCTATGCTGATAGGAGGATGATTTTCTACATGTTTA 1131	
DB		
QY	988 CAGGCTTGAGGAGCGCTCCATCTATGCTGATAGGAGGATGATTTTCTACATGTTTA 1047	
DB		

QY	1132 CTACTTCTTCCCTAAGAGAACCACTCCCTGATTTCTTCCGGCATCATAAAGCAGCTGC 1191	
DB		
QY	1048 CTACTTCTTCCCTAAGAGAACCACTCCCTGATTTCTTCCGGCATCATAAAGCAGCTGC 1107	
DB		
QY	1192 TCAGCTATTATATGAACCGAAGTGAAGTGTCTTAAATGCTTCCCTGCTTCCATATGA 1251	
DB		
QY	1108 TCAGCTATTATATGAACCGAAGTGAAGTGTCTTAAATGCTTCCCTGCTTCCATATGA 1167	
DB		
QY	1252 TTGCGAGGAGTTGTGAATCAGCCCTACCTTGTGTACTCGTTCACATGAAAGACCAAA 1311	
DB		
QY	1168 TTGCGAGGAGTTGTGAATCAGCCCTACCTTGTGTACTCGTTCACATGAAAGACCAAA 1227	
DB		
QY	1312 GCCATCCCTGTCCTCCAGCAAAACCCAGTCTCTGCTGGTGAATPCCACATCGCTATCTG 1371	
DB		
QY	1228 GCCATCCCTGTCCTCCAGCAAAACCCAGTCTCTGCTGGTGAATPCCACATCGCTATCTG 1287	
DB		
QY	1372 CAAAGACATTTCCATTCATTTTCAATGTTGACAAAGATATGACAAATTCGCAATTTGGCAA 1431	
DB		
QY	1288 CAAAGACATTTCCATTCATTTTCAATGTTGACAAAGATATGACAAATTCGCAATTTGGCAA 1347	
DB		
QY	1432 TGGCATCAGAAGGCTGATGAACAGAGAGACTTTTCAAGGAAAGCCCTAAATTTTCAAGAATA 1491	
DB		
QY	1348 TGGCATCAGAAGGCTGATGAACAGAGAGACTTTTCAAGGAAAGCCCTAAATTTTCAAGAATA 1407	
DB		
QY	1492 CTTTGAATTTCTGACTTCCAAAATCAACAGAGCTTTTAGCGGATCATGACTATGTTGAA 1551	
DB		
QY	1408 CTTTGAATTTCTGACTTCCAAAATCAACAGAGCTTTTAGCGGATCATGACTATGTTGAA 1467	
DB		
QY	1552 TATGCAAGTTTGTGTACGAGTGAAGAGTGGGAACTCTGTGGAAGAAATCTTCAAGGGT 1611	
DB		
QY	1468 TATGCAAGTTTGTGTACGAGTGAAGAGTGGGAACTCTGTGGAAGAAATCTTCAAGGGT 1527	
DB		
QY	1612 TATGCAAGTTTGTGTACGAGTGAAGAGTGGGAACTCTGTGGAAGAAATCTTCAAGGGT 1671	
DB		
QY	1528 TATGCAAGTTTGTGTACGAGTGAAGAGTGGGAACTCTGTGGAAGAAATCTTCAAGGGT 1587	
DB		
QY	1672 GTACCCCTGTGTGACAGATTTAGAAGATTTTACAGCAGAGGCTCTACCTCTCAGACAT 1731	
DB		
QY	1588 GTACCCCTGTGTGACAGATTTAGAAGATTTTACAGCAGAGGCTCTACCTCTCAGACAT 1647	
DB		
QY	1732 CCCAATTCACATGCACTGAGGATGTGGTCTTAAATAGGGAACAGCCGAGCTCAGA 1791	
DB		
QY	1648 CCCAATTCACATGCACTGAGGATGTGGTCTTAAATAGGGAACAGCCGAGCTCAGA 1707	
DB		
QY	1792 TGGCTTGAAGAGGCTGGGAAAGCTGAAGGCTACCTTTGAGCAAGCCCAAGCCCT 1851	
DB		
QY	1708 TGGCTTGAAGAGGCTGGGAAAGCTGAAGGCTACCTTTGAGCAAGCCCAAGCCCT 1767	
DB		
QY	1852 GGAGGAGGAGAGAAAAGACAGTACCTTCTGTGCTCCATATTTCCCTGTGAGGTGC 1911	
DB		
QY	1768 GGAGGAGGAGAGAAAAGACAGTACCTTCTGTGCTCCATATTTCCCTGTGAGGTGC 1827	
DB		
QY	1912 TCAGCAGCTGTGGCAAGGCAAGTCTTCAAGCCCAAGAGTTCAGTAATGTCCACCATGCT 1971	
DB		
QY	1828 TCAGCAGCTGTGGCAAGGCAAGTCTTCAAGCCCAAGAGTTCAGTAATGTCCACCATGCT 1887	
DB		
QY	1972 CTTCTCAGACATCGTTGGGTTCACTGCCATCTCTCCAGTCTCACCGCTCAGGTCAT 2031	
DB		
QY	1888 CTTCTCAGACATCGTTGGGTTCACTGCCATCTCTCCAGTCTCACCGCTCAGGTCAT 1947	
DB		
QY	2032 CACCATGCTCAATGCACTGTACACTCGTTCGACACAGAGTGTGGAGAGCTGGATGCTTA 2091	
DB		
QY	1948 CACCATGCTCAATGCACTGTACACTCGTTCGACACAGAGTGTGGAGAGCTGGATGCTTA 2007	
DB		
QY	2092 CAAAGTGGAGACCATTTGGCGATGCCCTATTGTGTAGCTGGGGATTTACACAAAGAGTGA 2151	
DB		
QY	2008 CAAAGTGGAGACCATTTGGCGATGCCCTATTGTGTAGCTGGGGATTTACACAAAGAGTGA 2067	
DB		
QY	2152 TACTCATGCTGTTCAGATAGCGCTGATGCCCTGAAGATGATGAGAGCTCTCTCATGAAGT 2211	
DB		
QY	2068 TACTCATGCTGTTCAGATAGCGCTGATGCCCTGAAGATGATGAGAGCTCTCTCATGAAGT 2127	
DB		

Qy	2212	TATGTCCTCCCATGGGAAACCTATCAAGATGCGAATTGGACTGCAGCTCTGGATCAGTTTT	2271
Db	2128		
		TATGTCCTCCCATGGGAAACCTATCAAGATGCGAATTGGACTGCAGCTCTGGATCAGTTTT	2187
Qy	2272	TGCTGGCGTCTGGGAGCTTAAATGCCCGCTACTGCTTTTGGAAACAATGCTCACTCT	2331
Db	2188		
		TGCTGGCGTCTGGGAGCTTAAATGCCCGCTACTGCTTTTGGAAACAATGCTCACTCT	2247
Qy	2332	GGCTAACAAATTTGAGTCCTGCAGTGTACCACGAAAAATCAATGTCAGGCCAACAACTTA	2391
Db	2248		
		GGCTAACAAATTTGAGTCCTGCAGTGTACCACGAAAAATCAATGTCAGGCCAACAACTTA	2307
Qy	2392	CAGATTACTCAAGACATGTCCTGGTTTCGTGTTTTACCCCTCGATCAAGGAGGAACCTCC	2451
Db	2308		
		CAGATTACTCAAGACATGTCCTGGTTTCGTGTTTTACCCCTCGATCAAGGAGGAACCTCC	2367
Qy	2452	ACCAAACTTCCCTAGTGAATCCCGGAAATCTGCCATTTTCTGGATGCTTTACCAACAAGG	2511
Db	2368		
		ACCAAACTTCCCTAGTGAATCCCGGAAATCTGCCATTTTCTGGATGCTTTACCAACAAGG	2427
Qy	2512	AACAAACTCAAAACCATGCTTCCAAAAGAAGATGTGGAAGATGGCAATGCCAATTTTTT	2571
Db	2428		
		AACAAACTCAAAACCATGCTTCCAAAAGAAGATGTGGAAGATGGCAATGCCAATTTTTT	2487
Qy	2572	AGGCAAGCATCAGGAATAGATTAGCAACCTATATACCTATTTATAGTCTTTGGGGTTT	2631
Db	2488		
		AGGCAAGCATCAGGAATAGATTAGCAACCTATATACCTATTTATAGTCTTTGGGGTTT	2547
Qy	2632	GACTCATTTGAAGATGTGTACAGCGCTCTGAAGCACATTAGGGATTTGTAGATGCTTAACAA	2691
Db	2548		
		GACTCATTTGAAGATGTGTAGAGCGCTCTGAAGCACATTAGGGATTTGTAGATGCTTAACAA	2607
Qy	2692	GCAGTATTAAAAATTTCAAGGACCAAGTCACAAATCTTCTCCTCTTTTAAACATGACAAA	2748
Db	2608		
		GCAGTATTAAAAATTTCAAGGACCAAGTCACAAATCTTCTCCTCTTTTAAACATGACAAA	2667
Qy	2749	-ATGTACTCTACTTCAGTACTTCAGCTCTTCAAGAAAAAATAAAAAA	2793
Db	2668	TATGTACTCTACTTCAGTACTTCAGCTCTTCAAGAAAAAATAAAAAA	2713

RESULT 9

AA197900/C	
ID	AA197900 standard; cDNA; 885 BP.
XX	
AC	AA197900;
XX	
DT	13-NOV-2001 (first entry)
XX	
DE	Human neuroblastoma expressed polynucleotide SEQ ID NO 3975.
XX	
KW	Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200166719-A1.
XX	
PD	13-SEP-2001.
XX	
XX	02-MAR-2001; 2001WO-JP01629.
XX	
PR	07-MAR-2000; 2000JP-0159195.
XX	
PA	(CHIB-) CHIBA PREFECTURE.
PA	(HISM) HISAMITSU PHARM CO LTD.
XX	
PI	Nakagawara A;
XX	
DR	WPI; 2001-565584/63.
XX	
PT	Nucleic acids originating in gene expressed in human neuroblastoma,
PT	useful as probe or primer in diagnosing prognosis of human
PT	neuroblastoma, malignancy and susceptibility indicator or tumour marker

RESULT 10

AAS94827

ID AAS94827 standard; DNA; 2954 BP.

XX

AC AAS9

XX
DT 14-FEB-2002 (first entry)

XX

DE Huma

XX
XZ

KW Huma

KW card

XX MY 01 PM '73

50
Hom

XX.

PN W020
YX

PN
WZU
YY

YY 18-0

FD
16-0
YY

XX DE 04-8

U4-A

XX 05-8

PR 05-A
vv

XX

PA (INC)

XX

PI Shift

PI Tai

XX

DR WPI;

XX

PT Comp

PT aggo

PT are

XX

PS claim

XX

CC The

CC sequ

cc diff

CC comp

comp thro

CC
CC

CC
poly
diag

cc
cc

cc
cc

0103	22
0103	22

2140	Db																			TGCCATGCTAAACCCCATTCGTCGTGATGGCCCTTGAAGATGATGGAACCTTTTCAGAAGAGGTG	2199
2213	QY																			ATGCTCTCCCATCGGAAACCTATCAAGATGCGAATTTGGACTGCACCTCTGGATCAGATTTTT	2272
2200	Db																			CTGACACCTTGATGGGAAGACCGATTACAGATGAGGATAGGAATTCACTCAGGCTCCGTCGTG	2259
2273	QY																			GCTGGCGTCGTTGGAGTTTAAAATGCCCCGTTACTGTCTTTTGGAAACAATGTCACCTCTG	2332
2260	Db																			GCTGGAGTTGTTGGGGTGGGAATGCCACGTTATTGCCTGTTTGGAAATATGTCACACATG	2319
2333	QY																			GCTTAACAACATTTTCAGTCCCTGACGTGTACACGAAAAATCAATGTCAGCCCAACAATTAC	2392
2320	Db																			GCAAGCAAAATTCAGTTCGGGAAGTCACCTTCGGCGCATCAATGTCAGCCCAACCACTTAC	2379
2393	QY																			AGATTACTCAAGACCTGCTCTGGTTTCGTGTTTACCCTTCGATCAAGGGAGGAACCTCCA	2452
2380	Db																			CAATTATTAACAGAGAAGAAAGTTTCACATTCATTCGCGGGTCTCGTGAAGAGCTTCCA	2439
2453	QY																			CCAACTCTCCCTAGTGAATCCCGGAAATCTGCCATTTTCTGTGATG	2498
2440	Db																			GACAACCTTTCCAAAGCAAAATTCCTGGGATCTGCTATTTCTCTGGAGG	2485

RESULT 11

RESULTS
AAL23886
ID AAL23886 standard; CDNA: 761 BP.

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 16343.

Human; breast cancer; cell marker; cytostatic; ss.

AA Homo sapiens. OS

XX PN WO200151628-A2.

AA PD 19-JUL-2001.

10-JAN-2001; 2001WO-US00798.

PR 14-JAN-2000; 2000US-0176077.

FR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.

FR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.

09-JUN-2000; 2000US-0211313.
25-JUL-2000; 2000US-0220534.

PA (MILL-) MILLENNIUM PREDICTIVE

PI · Lillie J, Xu Y, Wang Y, S

DR WPI; 2001-451856/48.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI · Lillie J, Xu Y, Wanq Y, Steinmann K;

AA
DR WPI; 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer

PS Claim 1; Page 2988-2989; 3695pp; English.

The invention relates to human breast cancer expressed polynucleotides (AAU07344-AAU26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity.

Sequence 761 BP; 234 A; 171 C; 202 G; 154 T; 0 other;
 SQ

Query Match	16.0%;	Score 482.2;	DB 22;	Length 761;
Best Local Similarity	95.24;	Pred. No. 1.4e-117;		
Matches 530;	Conservative 0;	Mismatches 23;	Indels 4;	Gaps 3;
Qy	412	ACATCCAGATTACCAAGTCTCCTTCAATTGATAGTGGCTCTGTTTGTCTAGCTCTCATATAA	471	
Db	209	ACATCCAGATTACCAAGTCTCCTTGAATGATAGTGGCTCTGTTTGTCTAGCTCTCATATAA	268	
Qy	472	GAATCTAGCTCATTCAGGAGGATCGCAGCAGGGTAAGAGACACCAACACCATGTTCTGT	531	
Db	269	GAATCTAGCTCATTCAGGAGGATCGCAGCAGGGTAAGAGACACCAACACCATGTTCTGT	328	
Qy	532	CACGAAGCTCAAGATCTCAAGATCATCAGAGAGTGTCTTTTCTCTTACTGGCACACAGG	591	
Db	329	CACGAAGCTCAAGATCTCAAGATCATCAGAGAGTGTCTTTTCTCTTACTGGCACACAGG	388	
Qy	592	TCAAGTTCCTTAACGAGCTCTTCAGAGGAGGAGGAGGCTTCAGAGAGCTGCAAAAGCAAC	651	
Db	389	TCAAGTTCCTTAACGAGCTCTTCAGAGGAGGAGGAGGCTTCAGAGAGCTGCAAAAGCAAC	448	
Qy	652	CGTGCCCATCTGTCAAGACATTCCTGAGAGAACATACAAGAAAGTCTTCTCTCAAGAAA	711	
Db	449	CGTGCCCATCTGTCAAGACATTCCTGAGAGAACATACAAGAAAGTCTTCTCTCAAGAAA	508	
Qy	712	AACCAGCTCGAGCGGAGTCTATCTTCACACTTTGGCAGAGAGTATTTGCAAACTGATTTT	771	
Db	509	AACCAGCTCGAGCGGAGTCTATCTTCACACTTTGGCAGAGAGTATTTGCAAACTGATTTT	568	
Qy	772	CCCAGAGTTTGAACGGCTGAATGTGACACTTCAGAGAACATTTGGCAAGACACAAATAAA	831	
Db	569	CCCAGAGTTTGAACGGCTGAATGTTCACACTTCAAGAAACATTTGCAAAAGCCCAAAATAA	628	
Qy	832	AGAAAGCAGGAATCTTTTGGAAAGAGAGACTTTTGAAAAACAATTCGACAGCAAGCAGT	891	
Db	629	AGAAAGCCGGAAATCTTTTGGAAAGAGAGACTTTTGAAAAAAC-ATTTCAGAACCAAGCAGT	687	
Qy	892	TGCAGCAGGAGTTCCTAGTGCAGGTTCATCAAGATCTCTTGGTCAAGAGGTTTTTAAAT	951	
Db	688	TGC-GCAGGAATTCCTAGTGCAGGTTTTT--AAAAATTTCTTTGGGGAAGGTTTTTAAAT	744	
Qy	952	ATGTTACGAGGAAGATG	968	
Db	745	ATGTTACGAGGAAG	761	

RESULT 12

ABL02451
ID ABL02451 standard; cDNA; 2357 BP.

XX
DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster expressed polynucleotide SEO ID NO 1835

AA Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene: ss.
KW

XX OS Drosophila melanogaster.

XX
PN WO200171042-A2.

XX
PD 27-SEP-2001.

XX
PF 23-MAR-2001: 2001WO-US09231.

XX
PR 23-MAR-2000: 2000US-191637P.

PR 11-JUL-2000; 2000S-0614150.
XX

PA (PEKE) PE CORP NY.
XX

PI Venter JC, Adams M, Li PWD, Myers EW;

[illegible]

XX WPI: 2001-656860/75.
DR P-PSDB; ABB58348.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 1835; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB857737-AB872072).
XX
XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2357 BP; 518 A; 688 C; 660 G; 491 T; 0 other;
XX
XX Query Match 10.5%; Score 315.6; DB 23; Length 2357;
XX Best Local Similarity 56.0%; Pred. No. 4.7e-73;
XX Matches 619; Conservative 0; Mismatches 484; Indels 3; Gaps 1;
XX
QY 1340 TCCTCGCTGGTATCCACATCGCTATTCTCAAGACATTTCCATTCTCATTTCTATGTTT 1399
DB 929 TCCGACCTGCAATGAACCTCGTCCAGCTTCTCGAAGATGTTTCCCTGGCACTTCATCG 988
QY 1400 GACAAGATATGACATTTCTCAATTTGGCAATGGCATCAGAGGCTGATGACAGGAGA 1459
DB 989 AACGAGCAGCTGGAATCGTGTGAGTTGGCGAGGATTCAGCAAAATGTACAAAGCCCTAC 1048
QY 1460 GACTTTCAAGGAAGCGCTAATTTTGAAGAACTACTTTGAAATCTGACTCCAA--AATC 1516
DB 1049 ATGGCGGACTTGGTTCAGCGGACCACTACTTTGATTTCAACAGCCCAAGGGATG 1108
QY 1517 AACGAGACGTTAGCGGGATCATGACTATGTTGAATATGACGTTTGTAGAGTGAGG 1576
DB 1109 ACCATGAAATTCGGGACATCTGAGGGCCACCTACAGCCCTTCTCTGATCGTTGAT 1168
QY 1577 AGATGGGACACTCTGTGAAGAAATCTCAAGGGTTATGGACCTCAAGGCCCAATGATC 1636
DB 1169 AATCCGCGGGCGCTGTGATTTCCCGGCATTGGACTGGATCAAGGGACAGATGGTG 1228
QY 1637 TACATTGTTGAATCCAGTCAATCTTGTGTTGGGTCAACCTGTGTGGACAGATTAGAA 1696
DB 1229 CACTGTCCGAGTCCAAATTCGTTGCTTCATTTGATTCCTTCCGACGCGCTGGAT 1288
QY 1697 GATTTTACAGGACGAGGCTCTACCTCTCAGACATCCCAATTCACAATGCATCGAGGAT 1756
DB 1289 GGGTTGACCTGCAACGAGCTCTTATCTCGGACATACCGCTGCAGACCCACCGAGGAG 1348
QY 1757 GTGGCTTTAATAGGGGAACAAGCCGAGCTCAAGATGGCTGAAGAAGAGGCTGGGGAAG 1816
DB 1349 GTGATCTGTGGGGCAACAGGCTCGGGCCAGGATGGACTCGCGCGGCGCATGGACAAG 1408
QY 1817 CTGAAGGCTACCTTTCAGCAAGCCCAACAGCCCTGGAGGAGGAGAAAGACAGTA 1876
DB 1409 ATCAAGAACAGATTAGGAGGCCCACTCGCGGGTGCACCAAGGAGGGAAGAACGTC 1468
QY 1877 GACCTTCTGTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAGCTGTGGCAAGGCAAGTT 1936
DB 1469 AGTCTGCTGACCTCATCTTTCGCGGAGATCGCGGAGAGCTTTGGCTGGCTCCTCC 1528
QY 1937 GTGCAAGCCCAAGATTTAGTAAATGTCACCATGCTTCTCAGACATCTGTGGTTCAT 1996
DB 1529 ATTGAGCCCAAGACGATATCCGATGTACGATCTCTTTCAGCGACATCTGTGCTTCACT 1588
QY 1997 GCCATCTGCTCCCAAGTGTACCGGCTGACGAGTCAATCACCATGCTCAATGCTATCACT 2056

DB 1589 AGCATCTGCTCGCGGCCACTCCCTTCATGTTGATCAGCATGCTGGAGGGCTGTACAAAG 1648
QY 2057 CCCTTCGACACAGCAGTGTGGAGAGCTGTCAAGGTGAGTGTCTCAAGGTGAGACCATTCGCGATGCC 2116
DB 1649 GACTTCGAGGAGTTTCTGCACTTCTTCGACGTGTACAGGTGAGAGCCATCGGGATGCC 1708
QY 2117 TATTGTGTAGCTGGGGGATTTACAAAGAGAGTGTACTCATCTGCTGTTTCAGATAGCGCTG 2176
DB 1709 TACTCGTGGCCAGTGGACTCCACGAGCGCTCCATCTACGAGCCCAAGGTTCGCTGG 1768
QY 2177 ATGGCCCTTGAACATGATGAGCTCTCTGATGAAGTTATGTTCTCCCATGAGAACCTATC 2236
DB 1769 ATGGCCCTTGAACATGATGAGCTCTCTGATGAAGTTATGTTCTCCCATGAGAACCTATC 1828
QY 2237 AAGATGCGAATTTGGACTGCACTCTGGATCAGTCTTGTGGGTCTCTGGAGTTAAATG 2296
DB 1829 AAGATGAGGATCGGCTGCAACGGCACTGTTTGGCGGAGTGTGGCGAGGAGATG 1888
QY 2297 CCGCTTACTGCTTTTGGAAACAATGCTACTCTGGGTAAACAATTTGAGTCTGCTGCA 2356
DB 1889 CCGAGTATTCCTCTCTCGGGCACAGCGTCAAGTCCCAAGGTTTCGAGTCGGGCAGC 1948
QY 2357 GTACCACGAAATCAATGTCAGCCCAACAACCTACAGATTACTCAAACTGCTCTGCT 2416
DB 1949 GAGGCGCTGAGATTAAATGTTAGCCCAACCAAGGATTCGCTGACCAAGCAGGAGG 2008
QY 2417 TTCTGCTTTTACCCCTCGATCAAGGGA 2442
DB 2009 TTTGAGTTTTCGCTGACGCGCGGGA 2034
XX
XX RESULT 13
XX ABL02450
XX ID ABL02450 standard; cDNA; 5259 BP.
XX
XX AC ABL02450;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 1832.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX PA (PEXE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
XX P-PSDB; ABB58347.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 1832; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5259 BP; 1324 A; 1334 C; 1321 G; 1280 T; 0 other;

Query Match 10.48; Score 314.8; DB 23; Length 5259;
Best Local Similarity 56.4%; Pred. No. 1.2e-72;
Matches 608; Conservative 0; Mismatches 467; Indels 3; Gaps 1;

QY 1340 TCCTCGCTGTGATCCACATGCTATTCTGCAAGACATTTCCATTCCATTCATGTTT 1399
Db TCCGACTGCAATGAACTCGTCAGCTTCTGCAAGATGTTCCCTGGCATTTCATCATG 2828
QY 1400 GACAAAGATAGCAATCTGCAATTTGGCAATGGCATGAGAGGCTGTGAACAGGAGA 1459
Db AACGAGCAGCTGGAACCTGGTGCAGTTGGCGAGGGGATTCAGCAAAATGTACAAGCCCTAC 2888
QY 1460 GACTTTCAAGGAAGCCATTTTGAAGATACTTTGAAATCTGACTCCAAA---AATC 1516
Db ATGCGGAGCTTTGGTTCGCCAGGCGACCACTACTTCTGATTTCAAACACCCAGGGATG 2948
QY 1517 AACACAGCTTTAGCGGGATCATGATATGTTGAATATGAGTGTGTTGACAGTGAGG 1576
Db ACCATGAATTCGGGACATCGTGAGCGGCACCTACAGCCTTTCCTGATCGGTTGAAT 3008
QY 1577 AGATGGACAACTCTGTGAGAAATCTTCAAGGGTTATGAGCCTCAAGGCCAAATGATC 1636
Db AATCCGCGCGCGCTGTGGATTTCCCGGCCATTGGACTGGAGATCAAGGACAGATGGT 3068
QY 1637 TACATTTGTAATCCAGTGAATCTGTTTGGGTACCCCTGTGGAGACAGATTAGAA 1696
Db CACTGCCCCAGTCCAAATCGTGCTCTTCATTTGGATCCCCCTTCCTGGAGCGGCTGGAT 3128
QY 1697 GATTTTACAGGACGAGGGCTACTACTCTCAGACATCCCAATTCACAAATGACATGAGGAT 1756
Db GGGTTGACCTGCAACGACCTCTTATCTCGGACATACCCGCTGCAGCAGCGCCACCGAGG 3188
QY 1757 GTGGTCTTATAGGGGACAAAGCCGAGCTCAGATGGCTGTGAAGAGGCTGGGGAAG 1816
Db GTATCTGTGGCGGACAAAGGCTCGGGCCAGGATGAGCTCGCGCGGCGCATGGACAAG 3248
QY 1817 CTGAAGGCTACCTTTGAGCAAGCCCAAGCCCTGGAGGAGGAGAGAAACAGACGTA 1876
Db ATCAAGACAGCATGAGGAGGCCAATCTCGGCGTGACCAAGGAGCGGAAGAGACGTC 3308
QY 1877 GACCTTCTGTGCTCCATTTTCCCTGTGAGGTTGCTCAGCAGCTGTGGCAAGGCAAGTT 1936
Db AGTCTGTCCACCTCATCTTTCCCGGGAGATCGCGAGAAGCTTGGCTGGGCTCCTCC 3368
QY 1937 GTCNAGCCAAAGATTCAGTATGTCACCATCTCTTTCACACATCGTTGGTTCACCT 1996
Db ATTGACCCAAAGACGTATCCCGATGTACGATCTCTGTTACGACACATCGTTGGCTTCACT 3428
QY 1997 GGCATCTGCTCCAGTCTCAGCGCTCAGGTCATACCATCTGCTCAATGACATGTACACT 2056
Db AGCATCTGCTCGGGGCCATCTCCCTTCATGTCATCAGATGCTGGAGGGGCTGTACAG 3488
QY 2057 CGCTTCGACAGGAGTGTGAGAGCTGGATGCTTACAAGGTGGAGACCAATTTGGCGATGCC 2116
Db GACTTCGACAGTTCCTCGACTTCTTCGAGCTGTACAAGGTGGAGACCAATCGGGGATGCC 3548
QY 2117 TATTGTGTACTGGGGATTACACAGAGAGTGTACTCATCTGCTTTCAGATAGCGCTG 2176
Db TACTGGCTGGCGAGTGGACCTCCACCGAGCCTTCCATCTACGACGCCCAACAGGTTGGCTGG 3608
QY 2177 ATGCCCCCTGAAGATGATGGAGCTCTCTGATGAAGTATGTCTCCCCATGGAGAACCTATC 2236
Db ATGCCCCCTGAAGATGATCGAGCGCTGCTCGAAGACACATTTACCCACGACGCGGAGCAATT 3668

QY 2237 AAGATGCAATTTGACTGCACTCTGGATCAGTTTTCGCGGCTGTTGGAGTTAAATG 2296
Db AATATGAGGATCGGCTGACACGGGACACTGTCTTGGCGGAGTGGTGGCAGGAATG 3728
QY 2297 CCCGTTACTGTCTTTTGGAAACAATGTCACTCTGGCTAACAAATTTGAGTCTCGAGT 2356
Db CCCAGGTATTGCTCTTTCGGGCACAGGTCACCATCGCCCAACAGTTTCGAGTCGGCAGC 3788
QY 2357 GTACCAGCAAAATCAATGTCAAGCCCAACACITACAGATTACTCAAGACTGTCCTG 2414
Db GAGCGCTGAAGATTATGTAGCCCAACCAAGGAGTAAGTGGTGTACTCTACTG 3846

RESULT 14
AAL15033
ID AAL15033 standard; cDNA; 649 BP.
XX
AC AAL15033;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 7490.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN W0200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
WP; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer
XX
PS Claim 1; Page 1347; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 649 BP; 194 A; 157 C; 166 G; 127 T; 5 other;

Query Match 10.2%; Score 308.4; DB 22; Length 649;
Best Local Similarity 96.0%; Pred. No. 1.9e-71;
Matches 389; Conservative 0; Mismatches 9; Indels 7; Gaps 7;

QY 412 ACATCCAGTTACCAGTGTCTTGAATGTAGTGGCTTCTGTTGTGCTCATATAA 471
Db ACATCCAGTTACCAGTGTCTTGAATGTAGTGGCTTCTGTTGTGCTCATATAA 268
QY 472 GAATACAGTCTATCAGGAGGAGATCGC-AGCAGGGTAAAGACACCAACCATGTTCT 530

|||||
Db 269 GAACACAGCTATCAGGAGGAGATCGCAAGCAGGGTAAGAGACACCAACCACTGTTCT 328
QY 531 GCACGAAGCTCAAGGATCTCAAGATCACAGGAGAGTGTCCTTTCTCTTACTGGCACCAG 590
|||||
Db 329 GCACGAAGCTCAAGGATCTCAAGATCACAGGAGAGTGTCCTTTCTCTTACTGGCACCAG 388
QY 591 GTCAGTCTCTAA-CGAGTCTTCAGAGGAGGAGCAGGAAGCTCAGAGAGC-TGCAAAAGC 648
Db 389 GTCAGTCTCTAAACCGAGTCTTCANAGAGGAGCAGGAAGCTCAGAGAGCTTGCAAAAGC 448
QY 649 AACCGTGCCTCATCTGTCACACATCTCTGAGAGACATACAGAAGTCTTCTCTCAAG 708
Db 449 AACCGTGCCTCATCTGTCACACATCTCTGAGAGACATACAGAAGTCTTCTCTTAAG 508
QY 709 AAAAAACAGT-CGAGCGCAGTCTATCTTCCACACTTTGGCA-GAGAGTATTTGCAAACTG 766
Db 509 AAAAAACAGTCCGAGCGAGTCTATCTTTACACTTTGGCANGAAGTATTTNCAAACTG 568
QY 767 ATTTTCCACAG-TTTGAACGGCTGAA-TGTTGCACCTTCAGAGAA 809
Db 569 ATTTTCCACAGTTTGAACGGCTGAATTGTTACACTTCANAAAA 613

RESULT 15

ABL04726

ID ABL04726 standard; cDNA; 3094 BP.

XX AC

XX XX

XX XX

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 8660.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX W0200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB60623.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 8660; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins

XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 3094 BP; 782 A; 727 C; 730 G; 855 T; 0 other;

Query Match 8.7%; Score 260.8; DB 23; Length 3094;

Best Local Similarity 61.2%; Pred. No. 2e-58;

Matches 421; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 1727 GACATCCCAATTCAAATGCACTGAGGAGTGTGCTCTTAATAGGGGAACAAGCCGAGCT 1786

Db 2 GACATACCGTGCACGACGCCACACGAGGAGTGATCTGTGTGGCGAACAAGCTTCGGGCC 61

QY 1787 CAAGATGGCTGAAGAAGAGCTGGGGAAGCTGAAGCTACCTTTGACGAGCCCAACAA 1846

Db 62 CAGATGGACTGCCCGCGCGCATGGACAAGATCAAGAACAGCATTTAGGAGGCCCACTCG 121

QY 1847 GCCCTGGAGGAGGAAGAAAAGACAGATGACCTTCTGTCTCCATATTTCCCTGTGAG 1906

Db 122 GCGGTGACCAAGGAGCGAAGAACGTCAGTCTGTGCACTCATCTTTCCCGCGGAG 181

QY 1907 GTTGCTCAGCAGCTGTGGCAAGGCAAGTTGTGCAAGCCAAAGATTTCAGTAATGTCAAC 1966

Db 182 ATCGCCGAGAAGCTTTTGGCTCCCTCCATTGACGCCAAGACGTATCCCGATGTTACG 241

QY 1967 ATGCTCTTCTCAGACATCGTTGGTTCACCTGCCATCTGCTCCAGTGTCTACCCCTGCG 2026

Db 242 ATCCTGTTCAGCGACATCGTTGGCTTCACTAGCACTGCTCGCGGCCCACTCCCTTATG 301

QY 2027 GTCATCACCATGCTCAATGCACCTACACTCGCTTCGACACAGCAGTGTGGAGAGCTGGAT 2086

Db 302 GTGATCAGCATGCTGGAGGGGCTGTACAAGGACTTCGACGAGTTCTGCGACTTCTTCGAC 361

QY 2087 GTCTACAAGGTGGAGACCATTGGCGATGCCATTGTTGTGTAGCTGGGGGATTACAAAGAG 2146

Db 362 GTGTACAAGGTGGAGACCATTGGGGATCGCTACTGCGTGGCCAGTGGCACTCCACCGAGCC 421

QY 2147 AGTGATACTCATGCTGTTTCAGATAGCGCTGATGGCCCTGAAGATGATGGAGCTCTCTGAT 2206

Db 422 TCCATCTACGACGCCACAGAGTTGCTGGATGGCCCTGAAGATGATGGAGCTCTGCTCG 481

QY 2207 GAAGTTATGTCTCCCCATGGAGAACCTATCAAGATGCGAATTTGGACTGCATCTGGATCA 2266

Db 482 AAGCACATTACCCACGACGCGAGCAAAATTAATGAGGATCGGCTGCACACGGGCACT 541

QY 2267 GTTTTGTGCTGCGCTGTTGGAGTTAAATGCCCCGTTACTGCTCTTTTGGAAACAATGTC 2326

Db 542 GTCTTGGCGGGAGTGGTGGGCAGGAAGATGCCAGGATTGCTCTTTTGGGCACAGCGTC 601

QY 2327 ACTCTGGCTAACAAATTTGAGTCTCCTGAGTGTACCAAGCAAAATTAATGTCAGCCCAACA 2386

Db 602 ACCATCCCAACAAGTTTCGAGTCGGGACGCGGCTGGAAGATTATGTTAGCCCAACC 661

QY 2387 ACTTACAGATTACTCAAAAGACTGTCTCTG 2414

Db 662 ACCAAGGAGTAAGTGGTACTCTACTG 689

Search completed: July 1, 2003, 13:26:14

Job time : 441.134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 12:30:42 ; Search time 2760.9 Seconds
(without alignments)
17686.068 Million cell updates/sec

Title: US-09-762-767A-1
Perfect score: 3015
Sequence: 1 cccctatggcattggcggg.....aaaaaaaaaaaaaaaaaaaaa 3015

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba.*
2: em_esthum.*
3: em_estmu.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1609.8	53.4	3054	11 AK004815	AK004815 Mus muscu
2	971.2	32.2	1169	13 BM544920	BM544920 AGENCOURT
3	882.4	29.3	884	14 BQ707076	BQ707076 AGENCOURT
4	738.2	24.5	850	9 AL543909	AL543909 AL543909
5	720.8	23.9	777	12 BG220276	BG220276 RST40049
6	707.4	23.5	829	9 AL563000	AL563000 AL563000

C	7	707	23.4	852	9	AL570204	AL570204
C	8	692.4	23.0	803	9	AI719730	AI719730
C	9	692.2	23.0	708	9	AA555318	AA555318
C	10	688.2	22.8	741	14	BQ74520	UI-H-E21-
C	11	687	22.8	1288	11	BC012627	Homo_sapi
C	12	673.4	22.3	824	13	BI771039	603055388
C	13	670	22.2	806	12	BG743485	602635511
C	14	655.2	21.7	767	9	AL527890	AL527890
C	15	654.8	21.5	658	14	BQ549434	1106h09.x
C	16	649.4	21.5	653	13	BM665903	UI-E-DX1-
C	17	625.2	20.7	1183	14	BM925915	AGENCOURT
C	18	624	20.7	777	12	BG203016	RST22385
C	19	620.2	20.6	716	9	AA779213	2353605.s
C	20	615	20.4	793	13	BI551700	603197377
C	21	605.4	20.1	964	13	BM469193	AGENCOURT
C	22	593	19.7	600	14	BQ581283	1106h09.y
C	23	567.4	18.8	595	10	AW513686	xc87f09.x
C	24	554.4	18.4	824	13	BI769310	603054578
C	25	551.8	18.3	644	10	AW006446	x
C	26	531.4	17.6	638	12	BG169925	602321619
C	27	506.2	16.8	631	10	AW951558	AW951558
C	28	505.4	16.8	539	9	AA189150	2q44g01.s
C	29	504.8	16.7	530	9	AL120503	DRFZP761N
C	30	495.2	16.4	734	13	BI551777	603197489
C	31	486.4	16.1	583	10	AW973582	EST385682
C	32	482.8	16.0	487	9	AL123854	q874f08.x
C	33	481	16.0	497	9	AL120382	DRFZP761D
C	34	478	15.9	1068	12	BG294057	602390917
C	35	477.4	15.8	583	9	AA193614	zr42g01.r
C	36	449.4	14.9	482	9	AA988462	or77r08.9
C	37	448.6	14.9	727	13	BI730880	603350511
C	38	442.2	14.7	705	9	AL598467	DRFZP3131
C	39	436.4	14.5	720	13	BI602173	603246294
C	40	435	14.4	911	12	BF679959	602154663
C	41	427.6	14.2	746	9	AL601127	DRFZP313N
C	42	424.8	14.1	447	13	BM662859	UI-E-CKO-
C	43	419	13.9	533	13	BM146581	TCAPPIE2
C	44	406.8	13.5	446	9	AA259214	nc17el0.i
C	45	388.6	12.9	675	13	BI546815	603189713

ALIGNMENTS

RESULT 1	AK004815	AK004815	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
LOCUS	AK004815	AK004815	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
DEFINITION	AK004815	AK004815	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
ACCESSION	AK004815	AK004815	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
VERSION	AK004815.1	AK004815.1	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
KEYWORDS	HTC; CAP trapper.	HTC; CAP trapper.	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
SOURCE	Mus musculus (strain:C57BL/6J) adult male lung cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library	Mus musculus (strain:C57BL/6J) adult male lung cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
ORGANISM	Mus musculus	Mus musculus	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
AUTHORS	Carninci, P. and Hayashizaki, Y.	Carninci, P. and Hayashizaki, Y.	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
TITLE	High-efficiency full-length cDNA cloning	High-efficiency full-length cDNA cloning	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	Meth. Enzymol. 303, 19-44 (1999)	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
MEDLINE	99279253	99279253	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
PUBMED	10349636	10349636	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
AUTHORS	Carninci, P. and Hayashizaki, Y.	Carninci, P. and Hayashizaki, Y.	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	Genome Res. 10 (10), 1617-1630 (2000)	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
MEDLINE	20499374	20499374	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002
PUBMED	11042159	11042159	3054 bp	mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200016007:guanylate cyclase 1, soluble, alpha 3, full insert sequence.	HTC 19-JAN-2002

Dd	441	GCACAGAGCTCTACCTGCGACACCCCTGGCAGAGAGTATTTTGCAGAGCTCATCTTCCAGAGT	500
Qy	780	TTCAAGCGGTGAATCTTGCACTTTCAGAGAAACATTTGGCAAGCACAAAATAAAGAAAGCA	839
Dd	501	CTGAGCGACTGAACCTTGCACTTCAGAGAACTTTGGCAAGCATATAATAGAAAGCA	560
Qy	840	GGAAATCTTTGGAAGAGAGACTTTTGAATAAACAATTTGCAGAGCAAGCTTTGCAGCAG	899
Dd	561	GGAAATCTTCAGAAAGAAAGACCTTTGAAATAAATCGCAGAAAGCAATTTGCAGCAG	620
Qy	900	GAGTTCACGTGAGGCTTATCAAGAAATCTCTTGTTGAAGAGGTTTAAATAATAGTTACG	959
Dd	621	GTGCCCCAGTGGAGGCGCTCAAGACTCTCTGGGGAGGAGCTGTTCAAGATCTGCTATG	680
Qy	960	AGGAAGATGAAACATCTCTTGGGGTGGTGGAGGCACCTTTAAAGATTTTAAACAGCT	1019
Dd	681	AGGAAGATGAGCACATTTTGGGCGTGGTGGGGCACCCCTGAAGACTTTCTAATAGCT	740
Qy	1020	TCAGTACCCCTTCGAACAGAGCAGCCATTTGCCAAGAAAGAGAAAGGGGAGGCTTG	1079
Dd	741	TCAGCACGCTCTCAAGCAGAGCAGCCACTGCAAGAGGGGAGAGGGCGGGGAGCAGCTG	800
Qy	1080	AGGACSCCTCCATCTATGSCCTGGATAGAGGAGTATTTCTACATGTTTACTACTTCT	1139
Dd	801	AGATGCCCTCCATCTATGCTCGGCAAGGACAGGACTTTCTAATGTTTACTACTTCT	860
Qy	1140	TCCCTAAGAGAACCCCTCCCTGATTTCTCCGGCATCATAAAGGCAGCTGCTCACGAT	1199
Dd	861	TCCGGAAGAGAACCCAGCCCTGCTTCTCCCTGGTATCATTAAGGGGCTGCTCGCATAC	920
Qy	1200	TATATGAACGGAAGTGAAGTGCCTTAATGCTCTCCCTGCTTCCATTAATGATTCAGCG	1259
Dd	921	TGTACGAAAGCCAGCTGGAGGTGCTCCCTGATGCTCTCCCTGCTTCCGAAAGTCACTGTACC	980
Qy	1260	AGTTTCTGAATCAGCCCTACTTGTGTTGCTACTCCGTTACATCAAAAGCACCAGCCATCCC	1319
Dd	981	AGTTTGTGAACCGCCCTATTTGCTCTACTCCGTTTCATGTGAAGAGCACCAGCCGCTCC	1040
Qy	1320	TGTCCCCAGCAAAACCCAGTCCCTCGCTGGTGAATTCACCACATCGCTATTCGCAAGACAT	1379
Dd	1041	TGTCCCCAGCAAGCCCGAGTCTCGCTAGTATGCCCGCTTGCTCTTCTGCAAGACTT	1100
Qy	1380	TTCCATTCATTTGTTGCAAGATATGACAAATTCGAATTTGGCAATGGCATCA	1439
Dd	1101	TCGCGTTCCATTTGATGTGGACCGAGACCTTGGCCATCTCGAGCTGGGTAAACGGCATCA	1160
Qy	1440	GAGGCTGATGAACGAGAGACTTTCAGGAAAGCCCTAAATTTCAAGAAATACATTGAAA	1499
Dd	1161	GAGGCTGTGAACAGAGGAGTTCACAGGAGCCCACTTTGAAGAGTCTTTGAAA	1220
Qy	1500	TTCTGACTCCAAAATCAACAGAGCTTTAGCGGATCATGACTATGTTGAATGCGAT	1559
Dd	1221	TTCTAACTCCAAAATCAACAGACATTTAGTGGATCATGACAAATGTTGAACATGCGAT	1280
Qy	1560	TGTTTGTAGCTGAGGAGATGGGACAACTCTGTGAAGAAATCTTCAAGGTTTATGACC	1619
Dd	1281	TTGTCTACCGGGTGAGGAGATGGGATAACTCGGTGAAGAAATCGTCAAGGTTTATGGATC	1340
Qy	1620	TCAAAGGCCAAATGATCATATGTTGAATCCAGTGCATCTGTTTGGGGTCACCCCT	1679
Dd	1341	TCAAAGGTCAAATGATCTACATCGTTGAATCCAGTGCCATCTGTCTCTTAGGGTCACCAT	1400
Qy	1680	GTGTGACAGATTAAGATTTTACAGGACGAGGCTCTACTCTCTCAGACATCCCAATTC	1739
Dd	1401	GTGTGACAGGCTGGAGATTTTACAGGACGGGGCTCTATCTGTCCGACATCCCAATTC	1460
Qy	1740	ACAATGCATGAGGATGTGCTTTAATAGGGGAACAAGCCCGAGCTCAAGATGGCCCTGA	1799
Dd	1461	ATAACCCCTGAGGGATGTGCTTTGATAGGGGAGCAGGCAAGGCTCAAGATGGCCCTCA	1520
Qy	1800	AGAAGAGGCTGGGAGCTGAAGGCTACCTTTGACCAAGCCCAACGACCCCTGGAGGAGG	1859
Dd	1521	AGAAGAGGTTGGGAGCTGAAGGCAACCTTGGAGCATGCCACCAAGCCCTGGAGGAGG	1580
Qy	1860	AGAAGAAAAGACAGCTAGACCTTCTGTCTCCATATTTCCCTGTGAGGTTGCTCAGCAGC	1919
Dd	1581	AGAAGAGAGGACAGTGGATCTGTCTCTATCTTCCCCTCTGAGGTTGCTCAGCAGC	1640
Qy	1920	TGTGGCAAGGGCAAGTTGTGCAAGCCAAAGATTCAAGTAAATGTCCACCATGCTTCTCTCAG	1979
Dd	1641	TGTGGCAAGGACAAAATTTGCAAGCCAAAGAAATTCAGCGAGGTCACCATGCTTCTCTCAG	1700
Qy	1980	ACATGCTTGGGTTCACTGCCATCTCTCCAGTCTCACCGCTCACAGGTCATCCATGTC	2039
Dd	1701	ATATCGTAGGGTTCACTGCTATCTCTCTCAGTGTTCACCTCTCTGAGGTCATCAGATGC	1760
Qy	2040	TCAATGCAGTGTACACTCGCTTCGACACGAGCTGTGGAGAGCTGGATCTCTACAAAGTGG	2099
Dd	1761	TCAACGCTCTCTACACTCGCTTTGACACGAGCTGTGGAGAGCTGGATGCTCTACAAAGTGG	1820
Qy	2100	AGACATNTGGCGATGCTTATTTGTGTAGCTGGGGGATTAACAACAAAGAGAGTGAATCATG	2159
Dd	1821	AGACCATCGGGATGCATATTTGTGTGGCAGGTGGATTCACAGAGAGAGTGCACCCATG	1880
Qy	2160	CTGTTTCAGATAGCTGATGGCCCTGAAGATGATGGAGCTCTCTGATGAAGTATGTCTC	2219
Dd	1881	CTGTCCAGATGACATGATGGCCCTGAAGATGATGGAGCTCTCCAATGAGGTCATGTCTC	1940
Qy	2220	CCCATGGAGAACCTATCAAGATGCGAATTTGGACTGCATCTCTGGATCAGTCTTTGCTGGCG	2279
Dd	1941	CCCACGGAGAACCTATCAAGATGCGAATTTGGACTACATCTCTGGATCAGTCTTTGCTGGAG	2000
Qy	2280	TGCTTGGAGTTAAATAGCCCGTTACTGTCTTTTTTGGAAACAATGTCTACTCTGGCTAACA	2339
Dd	2001	TTGTCGGAGTGAAGATGCCCGTATTTGCTGTTTGGAAACAATGTCTACTCTGGCTAACA	2060
Qy	2340	AAATTTGAGTCTGTCAGTGTACCACGAAAATCAATGTCCAGCCCAACACACTTACAGATTAC	2399
Dd	2061	AAATTTGAATCTCTGTCAGTGTGCTCGGAAATCAATGTCCAGCCCAACACACTACAGGTTAC	2120
Qy	2400	TCAAAGACTGTCTGCTTCTGCTTTTACCCCTCGCATCAAGGAGGAGAACTTCCACCAAACT	2459
Dd	2121	TCAAAGACTGTCTGCTTCTGCTTTTACCCCGGATCAAGGAGGAGGCTTCCACCAAACT	2180
Qy	2460	TCCCTAGTGAATCCCCGGAATCTGCCATTTTCTGGATGCTT---ACCAACAAGGAACAA	2516
Dd	2181	TCCCTAGTGACATTTCTGGGATCTGTCACTTTCTGGATGCTTATCACCATCAAGGACCTA	2240
Qy	2517	ACTCAAAACCATGCTTCCAAAGAAAGATGTGGAAGTGGCAATGCCAATTTTTTTAGGCA	2576
Dd	2241	ATTCCAACCATGGTTCAGGATAAGATGTGGAAGATGGAAACGCCAACCTTCTTAGGCA	2300
Qy	2577	AAGCATCAAGAAATAGATTAGCAACCTATATCTATTTAAGTCTTTTGGGTTTGACTC	2636
Dd	2301	AAGCGTCAGGATAGATTAGTCAGCCACAT-----GCTCTTATGTTTGTATGCC	2348
Qy	2637	ATTGAAGATGTGTAGAGCCTCTGAAAGCACTTTTAGGATGTGTAGATGCTAACAGCAGT	2696
Dd	23		

QY 2920 CCTGCAATCTCATATCCCTGGGGAATGCCATGGTATTATAAG-----TGTGTTTGTGATAG 2975
 Db 2645 TCACCTATGGCTGTGGCCACACAAAGCTGTGTATTATAAGTGTATGTGTGTGATAG 2704
 QY 2976 TGTGCTCAAAAAAATAAATAAATAAATAAATAAATAA 3008
 Db 2705 TGTGCTGTAAAAATCTCTGAATAGAAACAACA 2737

RESULT 2
 BM544920 1169 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT_6497474 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5588694
 DEFINITION 5', mRNA sequence.
 ACCESSION BM544920
 VERSION BM544920.1 GI:18776578
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1169)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM12360 Row: d Column: 07
 High quality sequence start: 39
 High quality sequence stop: 705.
 Location/Qualifiers
 1. 1169
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5588694"
 /clone_lib="NIH_MGC_125"
 /lab_host="DH10B"
 /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
 Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
 of three ovaries, from females ranging in age from 38 to
 49 yo. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 2.1 kb, insert size range 1-3.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 036."
 BASE COUNT 350 a 250 c 275 g 292 t 2 others
 ORIGIN

Query Match 32.2%; Score 971.2; DB 13; Length 1169;
 Best Local Similarity 95.5%; Pred. No. 8.8e-180;
 Matches 1074; Conservative 0; Mismatches 44; Indels 7; Gaps 7;

QY 812 TTGCAAGACACAAATAAAGAAAGACAGGAATCTTTGGAAGAGAACTTTGAAAAA 871
 Db 48 TTGCAAGACACAAATAAAGAAAGACAGGAATCTTTGGAAGAGAACTTTGAAAAA 107

QY 872 ACAATTCAGAGCAGTGGCAGCAGGAGTTCAGTGGAGGTTATCAAGATCTCTT 931
 Db 108 ACAATTCAGAGCAGCAGTGGCAGCAGGAGTTCAGTGGAGGTTATCAAGATCTCTT 167

QY 932 GGTGAAGAGGTTTTAAATATGTTACGAGGAGATGAAACATCTTTGGGGTGGTTGGA 991
 Db 168 GGTGAAGAGGTTTTAAATATGTTACGAGGAGATGAAACATCTTTGGGGTGGTTGGA 227

QY 992 GGCACCCCTAAAGATTTTTTAAACAGCTTCAGTACCCTTCGTAACACAGACGCCATTCG 1051

Db 228 GGCACCCCTAAAGATTTTTTAAACAGCTTCAGTACCCTTCGTAACACAGACGCCATTCG 287
 QY 1052 CAAGAAGCAGAAAAAGGGCGAGGCTTGAGGACGCCCTCCATCTTATGCCTGGATAAGAG 1111
 Db 288 CAAGAAGCAGAAAAAGGGCGAGGCTTGAGGACGCCCTCCATCTTATGCCTGGATAAGAG 347
 QY 1112 GATGATTTTCTACATGTTTACTACTTCTTCCCTTAAGAGAACACCTCCCTCATCTTCTCC 1171
 Db 348 GATGATTTTCTACATGTTTACTACTTCTTCCCTTAAGAGAACACCTCCCTCATCTTCTCC 407
 QY 1172 GGCATCAATAAGGAGCGTGTCTACGTATTATATGAAACGGAGTGGAGTGTCTGTTAATG 1231
 Db 408 GGCATCAATAAGGAGCGTGTCTACGTATTATATGAAACGGAGTGGAGTGTCTGTTAATG 467
 QY 1232 CCTCCCTGCTCCCATTAATGATTGAGCGAGTGTGTGAATCAGCCCTACTTGTGTACTCC 1291
 Db 468 CCTCCCTGCTCCCATTAATGATTGAGCGAGTGTGTGAATCAGCCCTACTTGTGTACTCC 527
 QY 1292 GTTCACATGAAAGCACCAGCCCATCCTGTCCCCAGCAAAACCCAGTCCCTCGCTGGTG 1351
 Db 528 GTTCACATGAAAGCACCAGCCCATCCTGTCCCCAGCAAAACCCAGTCCCTCGCTGGTG 587
 QY 1352 ATTCACACATCGCTTATCTGCAAGACATTTCCATTTCCATTTTCATTTTGACAAAAGATG 1411
 Db 588 ATTCACACATCGCTTATCTGCAAGACATTTCCATTTCCATTTTCATTTTGACAAAAGATG 647
 QY 1412 ACAATTCGCAATTTGGCAATGGCATGAGAGGCTGATGACAGAGAGAGACTTTCAAGGA 1471
 Db 648 ACAATTCGCAATTTGGCAATGGCATGAGAGGCTGATGACAGAGAGAGACTTTCAAGGA 707
 QY 1472 AAGCCTAATTTGAAGAACTACTTTGAAATCTGACTCCAAATAACAACAGACGCTTTAGC 1531
 Db 708 AAGCCTAATTTGAAGAACTACTTTGAAATCTGACTCCAAATAACAACAGACGCTTTAGC 767
 QY 1532 GGGATCATGACTATGTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1591
 Db 768 GGGATCATGACTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 827
 QY 1592 GTGAAGAAATCTTCAAGGTTATGAGCTCAAGGCAAAATGATCTACATTTGTTGAATCC 1651
 Db 828 GTGAAGAAATCTTCAAGGTTATGAGCTCAAGGCAAAATGATCTACATTTGTTGAATCC 887
 QY 1652 AGTCAATC-TGTGTTTGGGGTCACTGTTGGACAGATTTAGAAGA-TTTTACAGGAC 1709
 Db 888 AGTCAATC-TGTGTTTGGGGTCACTGTTGGACAGATTTAGAAGA-TTTTACAGGAC 947
 QY 1710 GAGGGCT-CTACCTCTCAGACATCCCAATTCACAAATGACTGAGGATGTTGTTAATA 1768
 Db 948 GAGGGCTCTACCTCTCAGAAATCCCAATTC-CCATGCACTGAAAGGATGTTGTTAATA 1006
 QY 1769 GGGGAACAAGCCGAGCTCAAGATGGCTGAAGAGAGGCTGGGGAAGCTGAAAGGCTTACC 1828
 Db 1007 GGGGAACAAGCCGAGCTTAAAT-GCCTGAAAAAAGCTTGGGAACCTTAAGG-TACC 1064
 QY 1829 CTTGAGCAAGCCCAAGCCCTGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1888
 Db 1065 CTTGAGCAAGCC-CCAGCCCCGGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1123
 QY 1889 TCATATTTCCCTGTGAGGTTGCTCAGCAGCTGTGGCAAGGCA 1933
 Db 1124 TCCAAATTTCCCTGGGAAGGTGCTCACCACCTGTGGCAGGGAA 1168

RESULT 3
 BQ707076
 LOCUS BQ707076
 DEFINITION AGENCOURT_8347306 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6279295
 5', mRNA sequence.
 ACCESSION BQ707076
 VERSION BQ707076.1 GI:21845975
 KEYWORDS EST.
 SOURCE human.


```
Db 1 CACTTTTCCTCGCTAGACGAGCAGCCTGGAAACAGACCCAGGCGGAGCACACCTGT 60
QY 187 GGGGAGGGAGCGCTCGAGAGCTTAGACACCCAGCCGGCGGTGATCTCACCATGTGC 246
Db 61 GGGGAGGGAGCGCTCGAGAGCTTAGACACCCAGCCGGCGGTGATCTCACCATGTGC 120
QY 247 GGATTTCCGAGCGCGCTCGAGAGCTTAGACACCCAGCCGGCGGTGATCTCACCATGTGC 306
Db 121 GGATTTCCGAGCGCGCTCGAGAGCTTAGACACCCAGCCGGCGGTGATCTCACCATGTGC 180
QY 307 GAAGCCACCAAGACTCGCGCTCTTGGAGAAAGCGTGACGAGGGGGCCACCGCGGTCTCCG 366
Db 181 GAASCACCAA----- 190
QY 367 CGCCTGTCTGCACCCCTGCTGAGCTGCTGACAGTACATGACATCCAGTTACCA 426
Db 191 -----ACATCCAGTTACCA 205
QY 427 GTGTCTTGAATGATAGTGGCTTCTGTTGTCTCAGTCTCATATAAGAACTACAGCTCATC 486
Db 206 GTGTCTTGAATGATAGTGGCTTCTGTTGTCTCAGTCTCATATAAGAACTACAGCTCATC 265
QY 487 AGGAGGAGATCGCAGCAGGCTTAAGACACCAACACCATGTTCTGCAGAGCTCAAGGA 546
Db 266 AGGAGGAGATCGCAGCAGGCTTAAGACACCAACACCATGTTCTGCAGAGCTCAAGGA 325
QY 547 TCTCAAGATCACAGGAGAGTGTCTTCTCTTACTGGCCAGGCTCAAGTCTCTCAAGGA 606
Db 326 TCTCAAGATCACAGGAGAGTGTCTTCTCTTACTGGCCAGGCTCAAGTCTCTCAAGGA 385
QY 607 GTCTTTCAGAGGAGCAGCAGGAGTCTGAGAGCTGCAAGACCAACCGTGCCTCATCTGCA 666
Db 386 GTCTTTCAGAGGAGCAGCAGGAGTCTGAGAGCTGCAAGACCAACCGTGCCTCATCTGCA 445
QY 667 AGACATTCCTGAGAGAGACATACAGAAAGTCTTCTCAAGAAACCAACAGTGGAGCCG 726
Db 446 AGACATTCCTGAGAGAGACATACAGAAAGTCTTCTCAAGAAACCAACAGTGGAGCCG 505
QY 727 AGTCTATCTTCACACTTTGGCAGAGAGTATTTGCAAACTGATTTTCCAGAGTTTGAACG 786
Db 506 AGTCTATCTTCACACTTTGGCAGAGAGTATTTGCAAACTGATTTTCCAGAGTTTGAACG 565
QY 787 GCTGAATGTGCACTTCAGAGACATTTGGCAAGCACAATAAAGAAAGCAGGAAATC 846
Db 566 GCTGAATGTGCACTTCAGAGACATTTGGCAAGCACAATAAAGAAAGCAGGAAATC 625
QY 847 TTTTGAAGAGAGACACTTTGAAACCAATTTGCAGAGCAAGCAGTTCGACGAGGATTC 906
Db 626 TTTTGAAGAGAGASACTTTGAAACCAATTTGCAGAGCAAGCAGTTCGACGAGGATTC 685
QY 907 AGTGGAGGTTATCAAGAAATCTCTTGTGAGAGGTTTAAATAATGTTTACGAGGAAGA 966
Db 686 AGTGGAGGTTATCAAGAAATCTCTTGTGAGAGGTTTAAATAATGTTTACGAGGAAGA 745
QY 967 TGAACATCCTTGGGGTGTGAGGACCCCTTAAAGATTTTAAACAGCTTCAGTAC 1026
Db 746 TGAACATCCTTGGGGTGTGAGGACCCCTTAAAGATTTTAAACAGCTTCAGTAC 805
QY 1027 CTTTCTCAACAGAGAGCCATTTGCCAAGAGCAGGAAAGGGG 1071
Db 806 CTTTCTCAACAGAGAGCCATTTGCCAAGAGCAGGAAAGGGG 850
```

```
RESULT 5
BG220276
LOCUS BG220276 777 bp mRNA linear EST 21-APR-2001
DEFINITION R64049 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG220276
VERSION BG220276.1 GI:13746297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 777)
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith,
E., Veloso, N., Kika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,
J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
2122/151
JOURNAL MEDLINE
COMMENT
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 501.
Location/Qualifiers
1. 777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT 239 a 165 c 168 g 205 t
ORIGIN
Query Match 23.9%; Score 720.8; DB 12; Length 777;
Best Local Similarity 97.5%; Pred. No. 8e-131;
Matches 753; Conservative 0; Mismatches 17; Indels 2; Gaps 2;
QY 839 AGGAATCTTTGGAAAGAGAGACTTTGAAAAACAATTCAGAGCAAGCAGTTGCAGCA 898
Db 4 AGGAATCTTTGGAAAGAGAGATCTTTGAAAAACAATTCAGAGCAAGCAGTTGCAGCA 63
QY 899 GGAGTTCAGTGGAGGTTATCAAGAATCTCTTGGTGAAGAGGTTTTTAAATATGTTAC 958
Db 64 GGAGTTCAGTGGAGGTTATCAAGAATCTCTTGGTGAAGAGGTTTTTAAATATGTTAC 123
QY 959 GAGAGATCAAAACATCTTGGTGGTGGAGGACCCCTTAAAGATTTTTTAAACAGC 1018
Db 124 GAGAGATCAAAACATCTTGGTGGTGGAGGACCCCTTAAAGATTTTTTAAACAGC 183
QY 1019 TTCAGTACCCTTCTGAAACAGAGCAGCATTGCCAAGACAGCAAGAAAGGGCAGGCTT 1078
Db 184 TTCAGTACCCTTCTGAAACAGAGCAGCATTGCCAAGACAGCAAGAAAGGGCAGGCTT 243
QY 1079 GAGAGCCCTCCATCTATGCTGGATAAGAGATGATTTCTACATGTTTACTACTTC 1138
Db 244 GAGAGCCCTCCATCTATGCTGGATAAGAGATGATTTCTACATGTTTACTACTTC 303
QY 1139 TTCCTTAAGAGAACCACTCCCTGATCTTCCCGGCATCATAAAGCAGCGCTCAGTA 1198
Db 304 TTCCTTAAGAGAACCACTCCCTGATCTTCCCGGCATCATAAAGCAGCGCTCAGTA 363
QY 1199 TTATATGAACGGAAGTGGAAAGTGTGTTAATGCCCTCCCTGCTTCCATATGATTGAGC 1258
Db 364 TTATATGAACGGAAGTGGAAAGTGTGTTAATGCCCTCCCTGCTTCCATATGATTGAGC 423
QY 1259 GAGTTGTGAATCAGCCCTTACTTGTGTACTCCGTTTCAATGAAAGACCAAGCCATCC 1318
Db 424 GAGTTGTGAATCAGCCCTTACTTGTGTACTCCGTTTCAATGAAAGACCAAGCCATCC 483
QY 1319 CTGTCCCGCAGCAACCCAGTCTCGTGTGATTCGCCACATCGCTATTTCTGCAAGACA 1378
Db 484 CTGTCCCGCAGCAACCCAGTCTCGTGTGATTCGCCACATCGCTATTTCTGCAAGACA 543
```



```

QY 1379 TTTCCATTCCATTTCATGTTTGACAAAGATATGACAATTTCTCAATTTGGCAATGGCATC 1438
Db 544 TTTCCATTCCATTTCATGTTTGACAAAGATATGACAATTTCTCAATTTGGCAATGGCATC 603
QY 1439 AGAAGGCTCATGAACAGGAGAGACTTTCAGGAAGAGCCCTAAATTTTCAAGAAATACCTTGAA 1498
Db 604 AGAAGGCTCATGAACAGGAGAGACTTTCAGGAAGAGCCCTAAATTTTGAAGAAATACCTTGAA 663
QY 1499 ATTCTGACTCCAAAATCAACAGAGAGCTTTAGCGGATCATGACTATGTTGAATATGCAG 1558
Db 664 ATTCTGACTCC-AAAATCAACAGAGAGCTTTAGC-GGAACATGACATTTCTTAATATGCCT 721
QY 1559 TTTGTTGACGAGTGAAGAGATGGGACAACTCTGTGAAGAAATCTTCAAGG 1610
Db 722 GTTGTGTACTAGTAGGAAGATGGGACCACTCTGTAAGAAATCTTCAAGG 773

RESULT 6
AL563000/c 829 bp mRNA linear EST 16-FEB-2001
LOCUS AL563000 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC027YN16 3
DEFINITION prime, mRNA sequence.
ACCESSION AL563000
VERSION AL563000.1 GI:12911980
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 829)
JOURNAL Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1 . 829
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC027YN16"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com
http://fulllength.invitrogen.com"

BASE COUNT 224 a 170 c 165 g 267 t 3 others
ORIGIN
Query Match 23.5%; Score 707.4; DB 9; Length 829;
Best Local Similarity 99.3%; Pred. No. 3.3e-128;
Matches 708; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 905 CCAGTGGAGGTATCAAGAATCTCTGTGGTGAAGAGCTTTTAAAAATATGTACGAGAA 964
Db 829 CCAGTGGAGGTATCAARAATCTCTGTGGTGAAGAGCTTTTAAAAATATGTACGAGAA 770
QY 965 GATGAAACATCTTCGGGTGGTGGAGGACCCCTTAAAGATTTTAAACAGCTTCAGT 1024
Db 769 GATGAAACATCTTCGGGTGGTGGAGGACCCCTTAAAGATTTTAAACAGCTTCAGT 710
QY 1025 ACCCTTCTGAACAGAGACCCATTGCCAAGAGAGAGAAAGGGCGAGCTTGAGGAC 1084

```

```

Db 709 ACCCTTCTGAACAGAGAGCCATTGCCAAGAGAGAGAAAGGGCGAGCTTGAGGAC 650
QY 1085 GCCTCCATTCTATGCTCGGATAGGAGGATGATTTTCTACATGTTTACTACTTCTTCCTT 1144
Db 649 GCCTCCATTCTATGCTCGGATAGGAGGATGATTTTCTACATGTTTACTACTTCTTCCTT 590
QY 1145 AAGAGAACACCTCCCTCGATTCTTCCCGGCATCATAAAGGCAGCTGCTCACGTATTATAT 1204
Db 589 AAGAGAACACCTCCCTCGATTCTTCCCGGCATCATAAAGGCAGCTGCTCACGTATTATAT 530
QY 1205 GAACGGAGTGGAGTGCCTTAATGCTCCCTGCTTCCATATGATTCAGCGAGTTT 1264
Db 529 GAACGGAGTGGAGTGCCTTAATGCTCCCTGCTTCCATATGATTCAGCGAGTTT 470
QY 1265 GTGAATCAGCCCTACTTGTGTACTCCGTTACATGAAAGACCAAGGCATCCCTGTCC 1324
Db 469 GTGAATCAGCCCTACTTGTGTACTCCGTTACATGAAAGACCAAGGCATCCCTGTCC 410
QY 1325 CCAGCAAAACCCAGTCTCGCTGGTGGATTCACATCGCTATTCGCAAGACATTTCCA 1384
Db 409 CCAGCAAAACCCAGTCTCGCTGGTGGATTCACATCGCTATTCGCAAGACATTTCCA 350
QY 1385 TTCCATTTTCATGTTTGACAAAGATATGACAATTTCTCAATTTGGCAATGGCATCAGAAG 1444
Db 349 TTCCATTTTCATGTTTGACAAAGATATGACAATTTCTCAATTTGGCAATGGCATCAGAAG 290
QY 1445 CTGATGAACAGAGAGACTTTCAGGAAAGCCCTAAATTTGAAGAAATCTTTGAAATCTG 1504
Db 289 CTGATGAACAGAGAGACTTTCAGGAAAGCCCTAAATTTGAAGAAATCTTTGAAATCTG 230
QY 1505 ACTCCAAAATCAACAGAGAGCTTTAGCGGATCATGACTATGTTGAATATGCAAGTTTGT 1564
Db 229 ACTCCAAAATCAACAGAGAGCTTTAGCGGATCATGACTATGTTGAATATGCAAGTTTGT 170
QY 1565 GTACGAGTGGAGATGGGACAACTCTGTGAAGAAATCTTCAAGGGTTATGGA 1617
Db 169 GTACGAGTGGAGATGGGACAACTCTGTGAAGAAATCTTCAAGGGTTATGGA 117

```

```

RESULT 7
AL570204/c 852 bp mRNA linear EST 16-FEB-2001
LOCUS AL570204 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI005YI13 3
DEFINITION prime, mRNA sequence.
ACCESSION AL570204
VERSION AL570204.1 GI:12926285
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 852)
JOURNAL Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1 . 852
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI005YI13"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/Note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,

```



```
QY 2400 TCAAGACATGCTCTGTTGTTGTTTACCCCTCGATCAAGGAGGAACTTCCACCAACT 2459
Db 385 TCAAGACATGCTCTGTTGTTGTTTACCCCTCGATCAAGGAGGAACTTCCACCAACT 326
QY 2460 TCCCTAGTCAAAATCCCGGAATCTGCATTTCTGGATGCTTACCAACAGGAACAACCT 2519
Db 325 TCCCTAGTCAAAATCCCGGAATCTGCATTTCTGGATGCTTACCAACAGGAACAACCT 266
QY 2520 CAAAACCATGCTTCCAAAAGAAAGATGTGGAAGATGCAATGCCAATTTTGTAGGCAAG 2579
Db 265 CAAAACCATGCTTCCAAAAGAAAGATGTGGAAGATGCAATGCCAATTTTGTAGGCAAG 206
QY 2580 CATCAGGAATAGATAGCAACCTATATACCTATTTAAGTCTTTGGGTTTGCATCAAT 2639
Db 205 CATCAGGAATAGATAGCAACCTATATACCTATTTAAGTCTTTGGGTTTGCATCAAT 146
QY 2640 GAAGATGCTAGAGCTCTGAAAGCACTTTAGGATTTAGATGCTTAAACATGACAAA 2755
Db 145 GAAGATGCTAGAGCTCTGAAAGCACTTTAGGATTTAGATGCTTAAACATGACAAA 26
QY 2700 AAAATTCAGGAGCCCAAGTACAAATCTTTCTCTGTTTAAACATGACAAA 2779
Db 85 AAAATTCAGGAGCCCAAGTACAAATCTTTCTCTGTTTAAACATGACAAA 26
QY 2756 CACTTCAGTACTCAGCTCTTCAA 2779
Db 25 CACTTCAGTACTCAGCTCTTCAA 2
```

RESULT 9

AA555318/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

708 bp mRNA linear EST 09-SEP-1997
nk79f06.s1 NCI_CGAP_Sch1 Homo sapiens cDNA clone IMAGE:1019747 3'
similar to gb:66534_cds1 GUANYLATE CYCLASE SOLUBLE, ALPHA-3 CHAIN
(HUMAN); mRNA sequence.

AA555318

AA555318.1

GI:2325857

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 708)

NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Stratagene, Inc., David B. Krizman,

Ph.D.

cDNA Library Arraying: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/dbbrp/image/image.html

Insert Length: 1229 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 480.

Location/Qualifiers

1..708

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1019747"

/clone_lib="NCI_CGAP_Sch1"

/tissue_type="Schwannoma tumor"

/lab_host="SOLR (kanamycin resistant)"

/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI

; Cloned unidirectionally. Primer: Oligo dt. Two pooled

bulk Schwannoma tumors. 5' adaptor sequence: 5'

FEATURES

source

RESULT 10

B0774520/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

741 bp mRNA linear EST 26-JUL-2002
UI-H-EZI-bcb-b-11-0-UI.s1 NCI_CGAP_Ch2 Homo sapiens cDNA clone
UI-H-EZI-bcb-b-11-0-UI 3', mRNA sequence.

B0774520

B0774520.1

GI:21982996

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 741)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE

**JOURNAL
COMMENT**

 $\frac{\partial y}{\partial z}$

FEATURES

.

BASE COUNT
ORIGIN

Query Match

**Best Local
Matches 7**

205
20y

db 74

211 dy

89 90

217 2y

629

223 QY

95 qb

229 QY

50 qb

235

44 Db

241 2y

38 qb

1

1

ORIGIN

Query Match 22.8%; Score 687; DB 11; Length 1288;
Best Local Similarity 98.7%; Pred. No. 3e-124;
Matches 704; Conservative 0; Mismatches 5; Indels 4; Gaps 1;

QY	2094	AGGTGAGACCATGGCGATGCTATTTGTAGCTGGGGATACAAAGAGAGTGATA	2153
Db	573	AGGTGAGACCATGGCGATGCTATTTGTAGCTGGGGATACAAAGAGAGTGATA	632
QY	2154	CTCATCTCTTCAGATAGCGCTGATGGCCCTGAAGATGATGGAGCTCTCTGATGAAGTGA	2213
Db	633	CTCATCTCTTCAGATAGCGCTGATGGCCCTGAAGATGATGGAGCTCTCTGATGAAGTGA	692
QY	2214	TGCTCTCCCATGAGAACCTATCAAGATGCGAATGCGATGCGACTCTGGATCAATGTTTG	2273
Db	693	TGCTCTCCCATGAGAACCTATCAAGATGCGAATGCGATGCGACTCTGGATCAATGTTTG	752
QY	2274	CTGGCTCTGTTGAGTAAATGCCCGTTACTGCTCTTTTGGAAACAATGCTACTCTGG	2333
Db	753	CTGGCTCTGTTGAGTAAATGCCCGTTACTGCTCTTTTGGAAACAATGCTACTCTGG	812
QY	2334	CTAACAAATTTGAGTCTGCTGAGTGTACACGAAATCAATGTCAGCCCAACAACCTTACA	2393
Db	813	CTAACAAATTTGAGTCTGCTGAGTGTACACGAAATCAATGTCAGCCCAACAACCTTACA	872
QY	2394	GATTACTCAAGACTGCTGCTGTTTCTGTTTACCCCTCGATCAAGGAGGAACTTCCAC	2453
Db	873	GATTACTCAAGACTGCTGCTGTTTCTGTTTACCCCTCGATCAAGGAGGAACTTCCAC	932
QY	2454	CAAACTCCCTAGTGAATCCCGGAATCTGCCATTTCTGGATGTTTACCACAGGAA	2513
Db	933	CAAACTCCCTAGTGAATCCCGGAATCTGCCATTTCTGGATGTTTACCACAGGAA	992
QY	2514	CAAACTCAAAACCATGCTTCCAAAGAAAGATGTGAAGATGCAATGCCAATTTTTTAG	2573
Db	993	CAAACTCAAAACCATGCTTCCAAAGAAAGATGTGAAGATGCAATGCCAATTTTTTAG	1052
QY	2574	GCAAGCATCAGGAATAGATAGCAACCTATATACCTATTTATAGCTTTTGGGGTTGA	2633
Db	1053	GCAAGCATCAGGAATAGATAGCAACCTATATACCTATTTATAGCTTTTGGGGTTGA	1112
QY	2634	CTCATTTAGATGTGTAGACCTCTGAAGCACTTTAGGATTTGATGGCTTAAACAGC	2693
Db	1113	CTCATTTAGATGTGTAGACCTCTGAAGCACTTTAGGATTTGATGGCTTAAACAGC	1172
QY	2694	AGTATTAAATTTTCAGGAGCAAGTCAATCTTCTCTGTTTAAACATGACAAA---A	2749
Db	1173	AGTATTAAATTTTCAGGAGCAAGTCAATCTTCTCTGTTTAAACATGACAAAATGTA	1232
QY	2750	TGTAATCACTTCAGTACTTCAAGTCTTCAAGAAAAAATAACCTTAAAAA	2802
Db	1233	TGTAATCACTTCAGTACTTCAAGTCTTCAAGAAAAAATAACCTTAAAAA	1285

RESULT 12

BI771039
LOCUS 60305388F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:520497 5',
DEFINITION mRNA sequence.
ACCESSION BI771039.1 GI:15762617
VERSION 1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 824)
NIH-MGC <http://mhc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnli.gov>
http://image.lnli.gov
High quality sequence stop: 776.

FEATURES

source
1. 824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:520497"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 250 a 189 c 216 g 169 t
ORIGIN

Query Match 22.8%; Score 673.4; DB 13; Length 824;
Best Local Similarity 88.8%; Pred. No. 1.5e-121;
Matches 814; Conservative 0; Mismatches 6; Indels 97; Gaps 4;

QY	176	AGGACACCTGTGGGGAGGAGCGCTGGAGGAGCTTAGAGACCCAGCGCGGTGATC	235
Db	1	AGGACACCTGTGGGGAGGAGCGCTGGAGGAGCTTAGAGACCCAGCGCGGTGATC	60
QY	236	TCACCATGTGCGGATTTGGAGGCGCCCTTGAGCTGTAGAGTCCGGAAGCACAGCC	295
Db	61	TCACCATGTGCGGATTTGGAGGCGCCCTTGAGCTGTAGAGTCCGGAAGCACAGCC	120
QY	296	CCGAGTGTGCGAAGCACCACCAAGACTGCGGCTCTTGGAAGAACGCTGAGCGGGCCAC	355
Db	121	CCGAGTGTGCGAAGCACCACCAAGACTGCGGCTCTTGGAAGAACGCTGAGCGGGCCAC	142
QY	356	CGGGGTCTCGCGCCTGTCTGACCCCTGTGCGCTGAGCTGCCTGACAGTGACATGACAT	415
Db	143	CGGGGTCTCGCGCCTGTCTGACCCCTGTGCGCTGAGCTGCCTGACAGTGACATGACAT	146
QY	416	CCAGTTACAGTGTCTTGAATTGATGAGTGGTCTTGTGTTGTCAGTCTCATATAAGAAC	475
Db	147	CCAGTTACAGTGTCTTGAATTGATGAGTGGTCTTGTGTTGTCAGTCTCATATAAGAAC	206
QY	476	TACAGTCTATCAGGAGGATCCGACGAGGTTAAGAGACACCAACCATGTTCTGCAGC	535
Db	207	TACAGTCTATCAGGAGGATCCGACGAGGTTAAGAGACACCAACCATGTTCTGCAGC	266
QY	536	AAGCTCAAGGATCTCAAGATCAGAGGAGTGTCTTCTCTTACTGGCACCAGGTCAC	595
Db	267	AAGCTCAAGGATCTCAAGATCAGAGGAGTGTCTTCTCTTACTGGCACCAGGTCAC	326
QY	596	GTTCTTAAAGAGTCTTTCAGAGGAGGAGGAGGAGGAGTGTCTTCTCTTACTGGCACCAGGTC	655
Db	327	GTTCTTAAAGAGTCTTTCAGAGGAGGAGGAGGAGGAGTGTCTTCTCTTACTGGCACCAGGTC	386
QY	656	CCATCTGTCAAGACATCTCTGAGAGAACATACAGAAAGTCTTCTCTCAAGAAACCC	715
Db	387	CCATCTGTCAAGACATCTCTGAGAGAACATACAGAAAGTCTTCTCTCAAGAAACCC	446
QY	716	AGTCGGAGCGGAGTCTATCTTACACTTTTGGCAGAGGATTTGCAAACTGATTTTCCCA	775
Db	447	AGTCGGAGCGGAGTCTATCTTACACTTTTGGCAGAGGATTTGCAAACTGATTTTCCCA	506

<pre>/db_xref="taxon:9606" /clone="CS0DCU27YN16" /clone_lib="ITI_NFL003_NBC3" /sex="male" /tissue_type="neuroblastoma cells" /lab_host="DH10B" /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"</pre>					
BASE COUNT	231 a	173 c	203 g	159 t	1 others
ORIGIN					
Query Match 21.7%; Score 655.2; DB 9; Length 767;					
Best Local Similarity 88.4%; Pred. No. 5.3e-118;					
Matches 762; Conservative 1; Mismatches 4; Indels 95; Gaps 1;					
QY	114	CAGAGTTTCCTACACTTTTCCTGCGCTAGAGCAGCGAGCGCTGGAACAGACCAGGC	173		
Dd	1	CAGAGTTTCCTACACTTTTCCTGCGCTAGAGCAGCGAGCGCTGGAACAGACCAGGC	60		
QY	174	GGAGGACACTGTGGGGGAGGAGCGCTTGAGGAGCTTAGAGCACCCAGCCGGCGGTGA	233		
Dd	61	GGAGGACACTGTGGGGGAGGAGCGCTTGAGGAGCTTAGAGCACCCAGCCGGCGGTGA	120		
QY	234	TCTCACCATGTGCGGATTTTCGAGGCGCGCCCTGGAGCTGTAGAGATCCGGAAGCACAG	293		
Dd	121	TCTCACCATGTGCGGATTTTCGAGGCGCGCCCTGGAGCTGTAGAGATCCGGAAGCACAG	180		
QY	294	CCCCGAGGTGTGGAAGCCACCAAGACTGCCGGCTCTTGGAAAGCGTGAGCAGGGGGCC	353		
Dd	181	CCCCGAGGTGTGGAAGCCACCA--	203		
QY	354	ACCGCGGTCTCCGCGCTGTCTGCACCCCTGTCGCCCTGAGCTGCCTGACAGTGACAATGAC	413		
Dd	204	-----AC	205		
QY	414	ATCCCAGTTACCAAGTGTCCTTGTAATTGATAGTGGCTTCGTTTTGCAGTCTCATATAAGA	473		
Dd	206	ATCCCAGTTACCAAGTGTCCTTGTAATTGATAGTGGCTTCGTTTTGCAGTCTCATATAAGA	265		
QY	474	ACTACAGCTCATCAGGAGGAGATCGCAGCAGGGTAAGAGACACCAACCATGTTCTGCA	533		
Dd	266	ACTACAGCTCATCAGGAGGAGATCGCAGCAGGGTAAGAGACACCAACCATGTTCTGCA	325		
QY	534	CGAAGCTCAAGGATCTCAAGATCACAGGAGAGTGTCCTTCTCTTACTGGCACCAGGTC	593		
Dd	326	CGAAGCTCAAGGATCTCAAGATCACAGGAGAGTGTCCTTCTCTTACTGGCACCAGGTC	385		
QY	594	AAGTTCTCAACGAGTCTTCAGAGAGGSCAGCAGGAAGCTCAGAGAGTGTCAAAAGCAACCG	653		
Dd	386	AAGTTCTCAACGAGTCTTCAGAGAGGSCAGCAGGAAGCTCAGAGAGTGTCAAAAGCAACCG	445		
QY	654	TGCCCATCTCTCAAGACATTCCTTGAGAGAACATACAGAAAGTCTTCTCTCAAGAAAAA	713		
Dd	446	TGCCCATCTCTCAAGACATTCCTTGAGAGAACATACAGAAAGTCTTCTCTCAAGAAAAA	505		
QY	714	CCAGTCGGAGCCGAGTCTATCTTCACACTTTGGCAGAGAGTATTTCGAACACTGTTTCC	773		
Dd	506	CCAGTCGGAGCCGAGTCTATCTTCACACTTTGGCAGAGAGTATTTCGAACACTGTTTCC	565		
QY	774	CAGAGTTTGAACGGCTGAATGTTGCATTCAGAGAACATTTGGCAAAGCAGCAAAATAAAG	833		
Dd	566	CAGAGTTTGAACGGCTGAATGTTGCATTCAGAGAACATTTGGCAAAGCAGCAAAATAAAG	625		
QY	834	AAAGCAGGAATCTTTGGAAAGAGAAGACTTTTGA AAAACAATTTGCAGAGCAAGCAGTTG	893		

Db	626	AAAGCAGGAAATCTTTGGAAAGAGAAGACTTTTGAAAAACAATTCGACAGCAAGCAGTTG	685
QY	894	CAGCAGGAGTTCAGTCGAGGTTATCAAGAAATCTCTTGCTGAAGAGGTTTTTAAATAT	953
Db	686	CAGCAGGAGTTCAGTCGAGGTTATCAAGAAATCTCTTGCTGRAGAGTTTTTAAATAT	745
QY	954	GTTACGAGGAAGATGAACAT	975
Db	746	GTTACGAGGAGATGAACAT	767
RESULT 15			
BQ549434/c			
LOCUS			
DEFINITION			
1106H09.x1 Human insulinoma Homo sapiens cdna clone IMAGE:6029417			
3' similar to SW:CYG3 HUMAN Q02108 GUANYLATE CYCLASE SOLUBLE,			
ALPHA-1 CHAIN ;, mRNA sequence.			
ACCESSION			
BQ549434			
VERSION			
Q0549434.1			
KEYWORDS			
EST.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 658)			
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,			
Lemishka, I., Secaere, M., Brestelli, J., Gradwohl, G., Clifton, S.,			
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blustein, A.,			
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas			
, M., Gibbons, M., McCann, R., Cole, R., Tsagarelshvili, R., Williams, T.,			
Jackson, Y., and Bowers, Y.			
Endocrine Pancreas Consortium			
Unpublished (2000)			
CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue			
Endocrine Pancreas Consortium			
Harvard University, Howard Hughes Medical Institute			
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,			
MA 02138			
Tel: 617-495-1812			
Fax: 617-495-8557			
Email: dmelton@biolhp.harvard.edu			
Library was constructed by Dr. J. J. Ferrer In vivo mass-excised to			
phagescript SK- by Dr. H. Inoue DNA sequencing by: Washington			
University Genome Sequencing Center For information on obtaining a			
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)			
Seq primer: -400P from Glibco			
High quality sequence stop: 485.			
Location/Qualifiers			
1. .658			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:6029417"			
/clone_lib="Human insulinoma"			
/tissue_type="insulinoma"			
/lab_host="Dr10B (phage-resistant)"			
/note="Organ: pancreas; Vector: phagescript SK-; Site:1:			
XhoI; Site:2: EcoRI; Constructed with lambda ZAPII system			
(Stratagene) by Dr. J. J. Ferrer, In vivo mass-excised to			
phagescript SK- by Dr. H. Inoue following the Washington			
University protocol			
(http://genome.wustl.edu/est/lambda_protocol.shtml).			
Please contact Hiroshi Inoue, MD/PhD for further			
information on this library (Metabolism Division, Permutt			
Laboratory, Washington University School of Medicine, Box			
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this			
is a Washington University Pancreas EST project library."			
BASE COUNT	181 a	131 c	140 g
ORIGIN			
Query Match	21.7%;	Score 654.8;	DB 14; Length 658;
Best Local Similarity	99.7%;	Pred. No. 6.5e-118;	
Matches 656; Conservative	0;	Mismatches 2;	Indels 0; Gaps 0;

```
Query Match      21.7%; Score 654.8; DB 14; Length 658;
Best Local Similarity 99.7%; Pred. No. 6.5e-118;
Matches 556; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 2135 TTACACAAAGAGATGATCTACTGCTGTTTCAGATAGCGTGTGATGGCCCTGAAGATGATG 2194
Db 658 TTACACAAAGAGAGTGTACTGCTGTTTCAGATAGCGTGTGATGGCCCTGAAGATGATG 599
Qy 2195 GAGCTCTCTGATGAAGTATCTCTCCCATGGAGAACCTATCAGATGCGAATTTGGACTG 2254
Db 598 GAGCTCTCTGATGAAGTATCTCTCCCATGGAGAACCTATCAGATGCGAATTTGGACTG 539
Qy 2255 CACTCTGGATCAGTTTTTGGCTGGCTGTTGAGTGAATGCCCCCTTACTGTCTTTTTT 2314
Db 538 CACTCTGGATCAGTTTTTGGCTGGCTGTTGAGTGAATGCCCCCTTACTGTCTTTTTT 479
Qy 2315 GGAACAATGTCACTCTGGCTTAACAAATTTGAGTGCCTGACAGTGTACACGAAAAATCAAT 2374
Db 478 GGAACAATGTCACTCTGGCTTAACAAATTTGAGTGCCTGACAGTGTACACGAAAAATCAAT 419
Qy 2375 GTCAGCCCCAACAACTTACAGATTACTCAAGACTGTCTGTTGCTGTTTACCCCTCGA 2434
Db 418 GTCAGCCCCAACAACTTACAGATTACTCAAGACTGTCTGTTGCTGTTTACCCCTCGA 359
Qy 2435 TCAAGGGAGGAACCTCCACCAAACTTCCCTAGTGAATCCCGGAATCTGCCATTTTCTG 2494
Db 358 TCAAGGGAGGAACCTCCACCAAACTTCCCTAGTGAATCCCGGAATCTGCCATTTTCTG 299
Qy 2495 GATGCTTACCAACAAAGGAACAACTCAAAACCATGCTTCCAAAAGAAAGATGTGGAAGAT 2554
Db 298 GATGCTTACCAACAAAGGAACAACTCAAAACCATGCTTCCAAAAGAAAGATGTGGAAGAT 239
Qy 2555 GGCAATGCCAATTTTTTAGGCAAGCATCAGGATAGATTAGCAACCTATATACCTATTT 2614
Db 238 GGCAATGCCAATTTTTTAGGCAAGCATCAGGATAGATTAGCAACCTATATACCTATTT 179
Qy 2615 ATAAGTCTTTGGGCTTTGACTCATTGAAGATGTGTAGAGCTCTGAAAGCACCTTTAGGGA 2674
Db 178 ATAAGTCTTTGGGCTTTGACTCATTGAAGATGTGTAGAGCTCTGAAAGCACCTTTAGGGA 119
Qy 2675 TTGTAGATGGCTAACAAAGCAGTATTAATAATTCAGGAGCCAAAGTCACAATCTTTCTCCTG 2734
Db 118 TTGTAGATGGCTAACAAAGCAGTATTAATAATTCAGGAGCCAAAGTCACAATCTTTCTCCTG 59
Qy 2735 TTTAACAATGACAAATCTACTCAGTCTCAGTCTCAGTCTTCAAGAAAAAATAAAAA 2792
Db 58 TTTAACAATGACAAATCTACTCAGTCTCAGTCTTCAAGAAAAAATAAAAA 1
```

Search completed: July 1, 2003, 17:26:01
Job time : 2768.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 13:06:07 ; Search time 101.089 Seconds
(without alignments)
9146.676 Million cell updates/sec

Title: US-09-762-767A-1
Perfect score: 3015
Sequence: 1 cccctatggcgtggcgggg.....aaaaaaaaaaaaaaaaaaaaa 3015

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2.6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2.6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2.6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2.6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2.6/ptodata/1/lna/PTUS_COMB.seq: *
6: /cgn2.6/ptodata/1/lna/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125.6	4.2	3784	1	US-07-623-033-1
2	121.4	4.0	3603	3	US-08-908-643C-85
3	121.4	4.0	3645	3	US-08-908-643C-83
4	121.4	4.0	3745	3	US-08-908-643C-84
5	121.4	4.0	3787	3	US-08-908-643C-82
6	91.4	3.0	4523	4	US-09-473-716-1
7	90.8	3.0	3518	4	US-09-412-210-2
8	86.6	2.9	3924	3	US-08-726-214-9
9	80.6	2.7	4008	3	US-08-307-896-5
10	80.6	2.7	4008	3	US-08-726-214-3
11	80.6	2.7	4008	3	PCT-US95-11808-5
12	77.8	2.6	3357	3	US-08-726-214-7
13	75.8	2.5	1652	3	US-08-726-214-17
14	75.8	2.5	3978	3	US-08-726-214-1
15	75.6	2.5	1812	4	US-09-008-097-3
16	75.6	2.5	3349	4	US-09-008-097-5
17	75.6	2.5	4942	4	US-09-474-076-1
18	74.4	2.5	4403765	4	US-09-103-840A-2
19	74.4	2.5	4411529	4	US-09-103-840A-1
20	72.8	2.4	4131	3	US-08-726-214-11
21	72.2	2.4	4046	1	US-07-793-961A-1
22	72.2	2.4	4046	1	US-08-240-357-1
23	71	2.4	4533	3	US-08-726-214-5
24	69.2	2.3	5199	3	US-08-726-214-13
25	62.4	2.1	4985	4	US-09-473-717-1
26	61.6	2.0	4473	3	US-08-894-173-1
27	61.6	2.0	4473	4	US-09-398-193-1

28	61	2.0	4601	3	US-08-726-214-15	Sequence 15, Appl
29	60.8	2.0	5515	4	US-09-398-193-98	Sequence 98, Appl
30	53	1.8	7218	1	US-08-232-463-14	Sequence 14, Appl
31	48.4	1.6	2092	3	US-08-307-896-6	Sequence 6, Appl
32	48.4	1.6	2092	5	PCT-US95-11808-6	Sequence 6, Appl
33	46.2	1.5	2320	4	US-09-813-817-1	Sequence 1, Appl
34	46.2	1.5	2320	4	US-09-978-197-1	Sequence 1, Appl
35	46	1.5	1652	4	US-09-627-650B-2	Sequence 2, Appl
36	46	1.5	1652	4	US-09-436-063C-2	Sequence 2, Appl
37	45.6	1.5	1143	4	US-09-149-476-119	Sequence 119, Appl
38	44.8	1.5	3459	2	US-08-980-060-3	Sequence 3, Appl
39	44.8	1.5	3459	4	US-09-307-185-3	Sequence 3, Appl
40	44.6	1.5	2028	2	US-09-211-930-12	Sequence 12, Appl
41	44.6	1.5	2028	3	US-09-340-993-12	Sequence 12, Appl
42	44.6	1.5	2028	4	US-09-468-442-12	Sequence 12, Appl
43	44.4	1.5	326	4	US-09-629-645A-17	Sequence 17, Appl
44	44.4	1.5	2920	3	US-09-276-400-1	Sequence 1, Appl
45	44.4	1.5	2920	4	US-09-448-076-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-623-033-1
; Sequence 1, Application US/07623033
; Patent No. 5237051
; GENERAL INFORMATION:
; APPLICANT: Garbers, David L.
; APPLICANT: Schulz, Stephanie
; TITLE OF INVENTION: CLONING THE ENTEROTOXIN RECEPTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGUMUS & CHESTNUT
; STREET: 100 South Wacker Drive - Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/623,033
; FILING DATE: 19901206
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: VU9018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 456-8000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3784 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; IMMEDIATE SOURCE:
; LIBRARY: gtl1 cDNA
; CLONE: #11 A
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..3235
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 11..76
; FEATURE:
; NAME/KEY: transit_peptide
; LOCATION: 1298..1369

FEATURE:
NAME/KEY: 5'clip
LOCATION: 1.10
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 3236...3784
US-07-623-033-1

Query Match 4.28; Score 125.6; DB 1; Length 3784;
Best Local Similarity 53.48; Pred. No. 6.4e-24;
Matches 313; Conservative 0; Mismatches 264; Indels 9; Gaps 2;

QY 1812 GGAAGCTGAAGGTACCTTTGAGCAAGCCACCACCAAGCCCTGGAGGAGGAGAGAAAGA 1871
DB 2334 GGAACCTGGAGCACCTGGTGGAGAAAGGACTCAGCTGTACAAAGCCGAGGACACAGG 2393
QY 1872 CAGTAGACCTCTGTGCTCCTATATTTCCCTGTGAGGTTGCTCAGCAGCTGTGGCAAGGC 1931
DB 2394 CTGACCACTTAATTTATGCTGCTCCACGGCTGGTGGTAAAGTCCCTGAAGGAGAAAG 2453
QY 1932 AAGTTGTGAAGCAAGAAAGTTCAGTAATGTCCACATGCTCTTCTCAGACATCGTTGGGT 1991
DB 2454 GCATCGTGGAGCAGAGCTGTACGAAGAAGTCAATCTATTTTCAGTGACATGTGCGGT 2513
QY 1992 TCAGTGCATCTGCTCCAGTGTCTACCGCTGCAGGTCTATCACCATGCTCAATGCACTGT 2051
DB 2514 TCAGGACCATCTGCAAGTACAGCAGCGCCATGGAGGTGGTGGACATGCTGAATGACATCT 2573
QY 2052 ACACCTCCCTTCGACCAAGCTGTGGAGAGCTGGATGTCTACAAGGTGGAGACCATGGCG 2111
DB 2574 ACAAGATTTTGACCAAGTGTGGATCACCACGCTCTACAAGGTAGAAACCATCGCG 2633
QY 2112 ATGCTATTGTGTAGCTGGGGGATTAC---ACAAAGAGAGTGATCTATCTGCTTTCAGA 2168
DB 2634 ATGCTAGCTGTGGCCAGCGCTGCTATGAGAACGCAACCGGCATGCACTGAGCA 2693
QY 2169 TAGCGTGTAGCGCTGAAGATGATGAGCTC-----TCTGATGAAGTTATGTCTCCC 2222
DB 2694 TTTTCAAGATGGCTTGGATCTCAGCTTCATGGGAGCTTTGAGCTGGAGCATCTCC 2753
QY 2223 ATGAGAACATATCAAGATCGAATTCGACTCTGCACTCTGATCAGTTTTCGCTGCGCTG 2282
DB 2754 CCGCCTCCCGTGTGATTCGATGGGTTTCTATCTGCCCCCTGTGCTGCTGCTGCTG 2813
QY 2283 TTGAGTTAAATCCCGCTTACTGTCTTTTGGAAACAATGTCATCTGCTGCTAACAAT 2342
DB 2814 TGGGATCAAGATCGCTGTTATGCTGTTTGGAGACATGTCAACACATGCTCCAGGA 2873
QY 2343 TTGAGTCTGCACTGTACCAAGAAATCAATGTACGCCCAACAAC 2388
DB 2874 TGGAGTCCACCGGCTTCCCTTAAGGATTCACATGAGCAGCTCCAC 2919

RESULT 2

US-08-908-643C-85
Sequence 85, Application US/080908643C
Patent No. 6120995

GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.
Pearlman, Joshua M.
Barber, Michael T.
Parkinson, Stephanie
Parkinson, Scott J.

TITLE OF INVENTION: COMPOSITIONS THAT SPECIFICALLY BIND TO
COLORECTAL CANCER CELLS AND METHODS OF
USING THE SAME

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6120995ris LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.

ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
FILING DATE: 07-Aug-1997
APPLICATION NUMBER: US/08/908,643C
CLASSIFICATION: N/A
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mark Deluca
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-2209
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 85:
SEQUENCE CHARACTERISTICS:
LENGTH: 3603
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 85:
US-08-908-643C-85

Query Match 4.08; Score 121.4; DB 3; Length 3603;
Best Local Similarity 53.28; Pred. No. 8.6e-23;
Matches 307; Conservative 0; Mismatches 261; Indels 9; Gaps 2;

QY 1812 GGAAGCTGAAGGTACCTTTGAGCAAGCCACCACCAAGCCCTGGAGGAGGAGAGAAAGA 1871
DB 2224 GAAACCTGGAGCAATCTGTTAGAGAAAGGACACAGCTGTACAAGGAGAGAGGACAGG 2283
QY 1872 CAGTAGACCTTGTGCTCCTATTTCCCTGTGAGGTTGCTCAGCAGCTGTGGCAAGGC 1931
DB 2284 CTCACAGACTTAATTTATGTTCTTCCAGGCTAGTGGTAAAGTCTCTGAAGGAGAAAG 2343
QY 1932 AAGTTGTGAGCAAGAGTTTCAGTAATGTCAACATGCTCTTCTCAGACATCGTTGGGT 1991
DB 2344 GCTTTGGAGCGGAACTATATGAGGAAGTTTCAATCTACTTCACTGACATGTAGGTT 2403
QY 1992 TCACCTGCTCTCTCCAGTCTCAGCGCTCAGGCTCATCAGCTCATCAGCTCTCAATGCACTGT 2051
DB 2404 TCACCTATCTCAATATCAGCAGCCCGCATGGAAGTGGTGGACATGCTTAATGACATCT 2463
QY 2052 ACACCTGCTTCGACCAAGCTGTGGAGAGCTGTCTACAAGGTGGAGACCATTTGGCG 2111
DB 2464 ATAAGAGTTTGGACCAATTTGATCATCATCATGCTGTACAAGTGGAAACCATCGGTG 2523
QY 2112 ATGCTATTGTGTAGCTGGGGATTAC---ACAAAGAGAGTATATCTATGCTCTCCC 2222
DB 2524 ATGCTATCATGTTGGCTAGTGTGCTTGCCTAAGAGAAATGGCAATCGCATGCAATAGACA 2583
QY 2169 TAGCGTGTAGCGCTGAAGATGATGAGCTC-----TCTGATGAAGTTATGCTCTCCC 2222
DB 2584 TTGCCAAGATGGCTTGGAAATCCTCAGCTTCATGGGGACCTTTGAGCTGGAGCATCTTC 2643
QY 2223 ATGAGAACCTATCAAGATGCGAATTTGGACTGTCACCTCTGGATCAGTTTTTGTGGGTG 2282
DB 2644 CTGGCTCCCAATATGATTCGATTTGGATTTGAGTTTCACTCTGCTGCTGCTGCTGAGTTG 2703
QY 2283 TTGAGTTAAATGCCCCGTTACTGTCTTTTGGAAACAATGTCATCTGCTGGCTTAACAAT 2342
DB 2704 TGGAAATCAAGATGCTGCTGTTATTTGATATTTGGAGATAGCTGCAACACAGCTCTAGGA 2763
QY 2343 TTGAGTCTGCACTGTACCAAGAAATCAATGCTCAG 2379
DB 2764 TGGAAATCAAGTCTGCTCTTTCAGATTTTCACGTGAG 2800

RESULT 3

US-08-908-643C-83
; Sequence 83, Application US/08908643C
; Patent No. 6120995

GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.
Pearlman, Joshua M.
Barber, Michael T.
Schultz, Stephanie
Parkinson, Scott J.

TITLE OF INVENTION: COMPOSITIONS THAT SPECIFICALLY BIND TO
COLORECTAL CANCER CELLS AND METHODS OF
USING THE SAME

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6120995rls LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/908,643C
FILING DATE: 07-Aug-1997
CLASSIFICATION: N/A

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: Mark Deluca

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJU-2209

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 83:

SEQUENCE CHARACTERISTICS:

LENGTH: 3645

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: unknown

SEQUENCE DESCRIPTION: SEQ ID NO: 83:

US-08-908-643C-83

Query Match 4.0%; Score 121.4; DB 3; Length 3645;

Best Local Similarity 53.2%; Pred. No. 8.6e-23;

Matches 307; Conservative 0; Mismatches 261; Indels 9; Gaps 2;

QY 1812 GGAAGCTGAAGCTACCTTGAGCAAGCCACCAGCCCTGGAGGAGGAGAGAAAAGA 1871
DB 2293 GAAAGCTGACATCTGTAGAGGAAAGACACAGCTGTACAGGACAGAGGGACAGG 2352
QY 1872 CAGTAGACTTCTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAGCTGTGGCAAGGCG 1931
DB 2353 CTGACAGACTTAACCTTTATGTGCTTCCAGGCTAGTGGTAAAGTCTCTGAGGAGAAAG 2412
QY 1932 AGTTGTGAGCCCAAGAGTTCAAGTATGTCACCATGCTCTTTCAGACATCGTTGGGT 1991
DB 2413 GCTTTGTGAGCGCGGAACATATATGAGGAAGTTACAATCTACTTCAGTGACATGTAGGTT 2472
QY 1992 TCACGTGACATCTGCTCCAGTCTCAGGCTGCTCAGGCTGATCAGCATGCTGCTGCTGCT 2051
DB 2473 TCACGTGACATCTGCAATATACACAGCCCATGGAAGTGGTGACATGCTTATGACATCT 2532
QY 2052 ACACGTGCTTGCACACAGCAGTGTGGAGAGCTGGATGCTGTACAAAGTGGAGACCATGGG 2111
DB 2533 ATAAGAGTTTGACCAACATGTTGATGATCATGATGCTGTACAAAGTGGAGAACCATCGGTG 2592

QY 2112 ATGCCTATGTGTAGCTGGGGGATTAC---ACAAAGAGAGTGTATCTATCTGCTGTTCAGA 2168
DB 2593 ATGCGTACATGTGTGCTAGTGTGCTTAAGAGAAATGGCAATCGCATGCAATAGACA 2652
QY 2169 TAGCGCTGATGGCCCTGAAGATGATGGAGCTC-----TCTGATGAAGTTATGCTCTCCC 2222
DB 2653 TTGCAAGATGGCCTTGGAAATCCCTCAGCTTCTGGGGACCTTTGAGCTGGAGCATCTTC 2712
QY 2223 ATGGAAGAACCTATCAAGATGCGAATGGACTGCACTCTGATCAGTCTTTTGTGCTGGCGTG 2282
DB 2713 CTGGCTCCCAATATGATTCGCATGGAGTTCACTCTGCTGCTGCTGCTGCTGAGTGG 2772
QY 2283 TTGGAGTTAAATGCCCGTTTACTGTCTTTTGGAAACAATGCTCACTCTGGCTAAACAAT 2342
DB 2773 TGGGAATCAAGATGCTGCTTATTGCTATTGAGATACGCTCAACACAGCCTCTAGCA 2832
QY 2343 TTGACTCTGCTGAGTGTACCAAGCAAAAATCAATGTCAG 2379
DB 2833 TGGAAATCAAGTGGCTCCCTTTGAGAAATTCACGTGAG 2869

RESULT 4

US-08-908-643C-84

; Sequence 84, Application US/08908643C

; Patent No. 6120995

GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.

Pearlman, Joshua M.

Barber, Michael T.

Schultz, Stephanie

Parkinson, Scott J.

TITLE OF INVENTION: COMPOSITIONS THAT SPECIFICALLY BIND TO
COLORECTAL CANCER CELLS AND METHODS OF
USING THE SAME

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6120995rls LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/908,643C

FILING DATE: 07-Aug-1997

CLASSIFICATION: N/A

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Mark Deluca

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJU-2209

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 84:

SEQUENCE CHARACTERISTICS:

LENGTH: 3745

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: unknown

SEQUENCE DESCRIPTION: SEQ ID NO: 84:

US-08-908-643C-84

Query Match

Best Local Similarity 4.0%; Score 121.4; DB 3; Length 3745;

Matches 307; Conservative 53.2%; Pred. No. 8.7e-23;


```
FILE REFERENCE: 44481-5027-01-US
CURRENT APPLICATION NUMBER: US/09/473,716
CURRENT FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: PCT/US98/13540
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/070,901
PRIOR FILING DATE: 1997-07-01
PRIOR APPLICATION NUMBER: 08/886,362
PRIOR FILING DATE: 1997-07-01
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 4523
TYPE: DNA
ORGANISM: human type V adenyllyl cyclase
FEATURE:
NAME/KEY: CDS
LOCATION: (139)..(3921)
US-09-473-716-1

Query Match      3.0%; Score 91.4; DB 4; Length 4523;
Best Local Similarity 49.9%; Pred. No. 1.2e-14;
Matches 230; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

QY 1945 CAAGAAGTTCAGTATGTCACCATGCTCTTCTCAGACATCGTTGGTTCACATGCGCATCTG 2004
Db 1524 CCAGAAACATGACAGAGGAGTCTCTTTGCTGACATCGAGGCTTACACAGCCTGGC 1583
QY 2005 CTCCAGTGTCTCAGGCTGACGATCATCACCATGCTCAATGCTCAATGCTGACATCGGCTTGA 2064
Db 1584 GTCCAGTGTCTCAGGCTGACGATCATCACCATGCTCAATGCTGACATCGGCTTGA 1643
QY 2065 CCAGCAGTGTGAGAGTGGATGCTACAAGGTGGAGACCATGTCGATGCTGCTATGCTG 2124
Db 1644 CAAGTGTGCTGAGAGTGGATGCTACAAGGTGGAGACCATGTCGATGCTGCTATGCTG 1703
QY 2125 AGCTGGGAGTGTACAAAGAGTGGATGCTACAAGGTGGAGACCATGTCGATGCTGCTATGCTG 2184
Db 1704 CTTCTGGGAGTGTACAAAGAGTGGATGCTACAAGGTGGAGACCATGTCGATGCTGCTATGCTG 1763
QY 2185 GAAGATGATGGAGTGTCTGATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2244
Db 1764 GGACATGATGGAGTGTCTGATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1823
QY 2245 AATTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2304
Db 1824 TGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1883
QY 2305 CTGCTCTTTTGGAAACAATGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2364
Db 1884 CGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1943
QY 2365 AAAATCAATGTCAGCCCAACCACTTACAGATTAATCTCAAG 2405
Db 1944 ACGCATCCATCATCACCAGGCTACACTCACTACCTGATG 1984

RESULT 7
US-09-412-210-2
; Sequence 2, Application US/09412210
; Patent No. 6403358
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 21529, A NOVEL ADENYLATE CYCLASE
; FILE REFERENCE: 5800-47
; CURRENT APPLICATION NUMBER: US/09/412,210
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3518
; TYPE: DNA
; ORGANISM: Homo sapiens

FILE REFERENCE: 44481-5027-01-US
CURRENT APPLICATION NUMBER: US/09/473,716
CURRENT FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: PCT/US98/13540
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/070,901
PRIOR FILING DATE: 1997-07-01
PRIOR APPLICATION NUMBER: 08/886,362
PRIOR FILING DATE: 1997-07-01
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 4523
TYPE: DNA
ORGANISM: human type V adenyllyl cyclase
FEATURE:
NAME/KEY: CDS
LOCATION: (139)..(3921)
US-09-473-716-1

Query Match      3.0%; Score 90.8; DB 4; Length 3518;
Best Local Similarity 50.5%; Pred. No. 1.6e-14;
Matches 221; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

QY 1941 AAGCCAAAGTTCAGTATGTCACCATGCTCTTCTCAGACATCGTTGGTTCACATGCGCA 2000
Db 1040 ATGTCAGAGGACCAAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1099
QY 2001 TCTGCTCCAGTGTCTCAGGCTGACGATCATCACCATGCTCAATGCTCAATGCTCAATGCTCA 2060
Db 1100 TGGCCAGGAGTGTCTCCCTTAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159
QY 2061 TGGCCAGGAGTGTCTCCCTTAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2120
Db 1160 TGGCCAGGAGTGTCTCCCTTAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1219
QY 2121 GTGTAGTGTGGGAGTGTACAAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2180
Db 1220 ACTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1279
QY 2181 CCTCAAGATGATGAGGCTCTCTGATGAAGTGTATGCTCCCATGAGAACCTATCAACA 2240
Db 1280 GCTTGGACATGTCGGGGCCATCAGGAACTGCGGGCAGCCTGCGGTGGACATCAACA 1339
QY 2241 TCGGAATGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2300
Db 1340 TCGGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1399
QY 2301 GTTACTGTCTTTTGGAAACAATGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2360
Db 1400 AGTACAGCTGTGGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1459
QY 2361 CACGAAAATCAATGTCA 2378
Db 1460 CAGGCGAGTGCACATCA 1477

RESULT 8
US-08-726-214-9
; Sequence 9, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/726,214
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
```

FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:450
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3924 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-726-214-9

Query Match 2.9%; Score 86.6; DB 3; Length 3924;
Best Local Similarity 49.2%; Pred. No. 2.3e-13;
Matches 227; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY 1945 CAAGAAGTTCAGTAATGTCACCATGCTCTTCTCAGACATCGTTGGGTTCACTGCCATCG 2004
DB 899 CCAGAAATGACAAATGTGACATCTCTTTGCTGACATCGAAGGCTTCCTAGCTGCG 958
QY 2005 CTCACAGTGTACCGCTGAGGTGATCATACCATGCTCAATGCTGACACTGACCTGCTCGA 2064
DB 959 ATCCCACTGTACTGCCCAAGAACTGCTCATGACCTCAACGAGCTCTTCCCGCGCTTGA 1018
QY 2065 CCAGCAGTGTGAGAGCTGATGCTACAAAGTGGAGACCATTTGGGAGTGGCTATTTGCT 2124
DB 1019 CAGTTGGCTGCGGAAATCACTGCTTACGGATTAAGATCTCTCGGGGATTTGTTACTAC 1078
QY 2125 AGCTGGGGGATTACAAAGAGAGTGTATCTATCTGTTTCAGATAGCGCTGATGGCCCT 2184
DB 1079 TGTCTCGGGCTGCTGGAAGCCAGAGCTGACACCGCCCTGCTGCTGAGAGATGGGAAT 1138
QY 2185 GAAGATGATGGAGCTCTCTGATGAAGTATGTTCTCCCATGAGAACCTTATCAAGATGG 2244
DB 1139 GGACATGATGAGGAGGCTCTGCTGCTGCGGGAGGTGACAGGGGTGAAGCTGAACATGG 1198
QY 2245 AATGGAGTGCATCTGATGATCTTTTGTGCTGCGCTGTTGGAGTTAAATGCCCGTTA 2304
DB 1199 TGTGGGAAATCAGAGGGGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1258
QY 2305 CTGCTCTTTTGGAAACAATGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2364
DB 1259 CGAGTGTGTCTAAACGATGCTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1318
QY 2365 AAAATCAATGTGAGCCCAACACTTACAGATTACTCAAG 2405
DB 1319 CCGATCCATCATCAAGGCCACACTCAACTCACTGAACG 1359

RESULT 9
US-08-307-896-5
Sequence 5, Application US/08307896C
Patent No. 6034071
GENERAL INFORMATION:
APPLICANT: Iyengar, Srinivas Ravi
TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND ADENYLYL
TITLE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS
FILE REFERENCE: 29770
CURRENT APPLICATION NUMBER: US/08/307,896C
CURRENT FILING DATE: 1994-09-16
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 4008
TYPE: DNA
ORGANISM: Rattus norvegicus
US-08-307-896-5

Query Match 2.7%; Score 80.6; DB 3; Length 4008;

Best Local Similarity 49.2%; Pred. No. 9.7e-12;
Matches 212; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 1931 CAAGTTGTGCAAGCCCAAGATTCAGTAATGTCACCATGCTCTTCTCAGACATCGTTGGG 1990
DB 901 CACAATCTGTATGTCAAAGCAGACACCAACGCTGAGCATATTATAGCTGACATTTTGGC 960
QY 1991 TTCACTGCCATCTGCTCCAGTGTCTCAGCTGCTGAGGTGATGATGATGATGATGATGAT 2050
DB 961 TTCACCCGCTTGCAGCGATGCTCCCTCGGAGCTGCTCCACATGCTGAATGAATC 1020
QY 2051 TACACTCGCTGCGACCGAGTGTGGAGAGTGTGATGATGATGATGATGATGATGATGATG 2110
DB 1021 TTTGGAAAGTTTGTATCAATGCAAGGAGAAATGAATGATGATGATGATGATGATGATGAT 1080
QY 2111 GATGCTTATGTTAGTGGGGGATTCACAAAGAGAGATGATGATGATGATGATGATGATGAT 2170
DB 1081 GACTGCTATGTTGTTTCCGGGCTCCCTATATGATGATGATGATGATGATGATGATGATGAT 1140
QY 2171 GCGCTGATGGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2230
DB 1141 GTGAAATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 2231 CCTATCAAGATGCGAATTTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2290
DB 1201 GATATCAACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 2291 AAAATGCCCCGTACTGCTTTTGGAAACAATGATGATGATGATGATGATGATGATGATGAT 2350
DB 1261 CAGAAGTGGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 2351 TGCAGTGTACC 2361
DB 1321 GGAGGAGTCCC 1331

RESULT 10
US-08-726-214-3
Sequence 3, Application US/08726214
Patent No. 6107076
GENERAL INFORMATION:
APPLICANT: Tang, Wei-Jen
APPLICANT: Gilman, Alfred G.
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,214
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 60/005,498
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:450
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 4008 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-726-214-3

Query Match 2.7%; Score 80.6; DB 3; Length 4008;

Best Local Similarity 49.2%; Pred. No. 9.7e-12;
Matches 212; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 1931 CAAGTTGTGCAAGCCCAAGAGTTCACTAATGTCACCATGCTCTTCACACATCGTTGGG 1990
DB 901 CACAATCTGTATGTCAACGACACACCAACGTCGAGCATATTATACCTGACATTTGTGGC 960
QY 1991 TTCACGCGCTTGCAAGCGATTGCTCCCTGGGAACTGGTCCACATGCTGAATGAATC 1020
DB 961 TTCACCGCGCTTGCAAGCGATTGCTCCCTGGGAACTGGTCCACATGCTGAATGAATC 1020
QY 2051 TACACTCGCTTCGACACGAGTGTGGAGCTGGATGCTTACAAGGTGGAGACCATTTGGC 2110
DB 1021 TTTGGGAAGTTTGATCAAAATAGCAAGGAGAATGAATGATGAGAAATTTAGGA 1080
QY 2111 GATGCGCTATGTAGTCTGCGGGAATACACAAAGAGTGTACTGCTGTTAGATA 2170
DB 1081 GACTGTATTAAGTGTCTTCCGGGCTCCCTATATCACTCCCTAACCATGCCAAGAACTGT 1140
QY 2171 GCGCTGATGCGCTTGAAGATGATGGAGCTCTCTGTATGAAGTGTATGCTCCCATGGAGAA 2230
DB 1141 GTGAAATGGATGGATATGTCGAAGCAATTAAGAAAGTGAAGGATGCTACCGGAGTT 1200
QY 2231 CCTATCAAGATGCGAATGGAGTGCATCTGTATGATGATGATTTTGTGCGGCTGTTGGAGTT 2290
DB 1201 GATATCAACATGCGTGTAGAGTGCATCTGGAAGCTTCTCTGTGTTGATTTGCTC 1260
QY 2291 AAAATCCCGTTACTGCTCTTTTGGAAACAATGTCACCTGCTGCTAACAATTTGAGTCC 2350
DB 1261 CAGAAGTGGCAGTATGATGTGGTCTCATGATGTTACTCTGCGCAACCATGGAAGCT 1320
QY 2351 TGCAGTGTACC 2361
DB 1321 GGAGGAGTCCC 1331

RESULT 11

PCT-US95-11808-5
Sequence 5, Application PC/TUS9511808
GENERAL INFORMATION:
APPLICANT: Iyengar, Srinivas Ravi V.
TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND
TITLE OF INVENTION: ADENYLIL
TITLE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue and
ADDRESSEE: Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11808
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/307,896
FILING DATE: 16-SEP-1994

ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S.
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: 29970 165/28755
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 408-2500
TELEFAX: (212) 765-2519
TELEX: 650 6111063

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 4008 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

PCT-US95-11808-5

Query Match 2.7%; Score 80.6; DB 5; Length 4008;

Best Local Similarity 49.2%; Pred. No. 9.7e-12;

Matches 212; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 1931 CAAGTTGTGCAAGCCCAAGAGTTCACTAATGTCACCATGCTCTTCACACATCGTTGGG 1990
DB 901 CACAATCTGTATGTCAACGACACACCAACGTCGAGCATATTATACCTGACATTTGTGGC 960
QY 1991 TTCACGCGCTTGCAAGCGATTGCTCCCTGGGAACTGGTCCACATGCTGAATGAATC 1020
DB 961 TTCACCGCGCTTGCAAGCGATTGCTCCCTGGGAACTGGTCCACATGCTGAATGAATC 1020
QY 2051 TACACTCGCTTCGACACGAGTGTGGAGCTGGATGCTTACAAGGTGGAGACCATTTGGC 2110
DB 1021 TTTGGGAAGTTTGATCAAAATAGCAAGGAGAATGAATGATGAGAAATTTAGGA 1080
QY 2111 GATGCGCTATGTAGTCTGCGGGAATACACAAAGAGTGTACTGCTGTTAGATA 2170
DB 1081 GACTGTATTAAGTGTCTTCCGGGCTCCCTATATCACTCCCTAACCATGCCAAGAACTGT 1140
QY 2171 GCGCTGATGCGCTTGAAGATGATGGAGCTCTCTGTATGAAGTGTATGCTCCCATGGAGAA 2230
DB 1141 GTGAAATGGATGGATATGTCGAAGCAATTAAGAAAGTGAAGGATGCTACCGGAGTT 1200
QY 2231 CCTATCAAGATGCGAATGGAGTGCATCTGTATGATGATTTTGTGCGGCTGTTGGAGTT 2290
DB 1201 GATATCAACATGCGTGTAGAGTGCATCTGGAAGCTTCTCTGTGTTGATTTGCTC 1260
QY 2291 AAAATCCCGTTACTGCTCTTTTGGAAACAATGTCACCTGCTGCTAACAATTTGAGTCC 2350
DB 1261 CAGAAGTGGCAGTATGATGTGGTCTCATGATGTTACTCTGCGCAACCATGGAAGCT 1320
QY 2351 TGCAGTGTACC 2361
DB 1321 GGAGGAGTCCC 1331

RESULT 12

US-08-726-214-7
Sequence 7, Application US/08726214
Patent No. 6107076
GENERAL INFORMATION:
APPLICANT: Tang, Wei-Jen
APPLICANT: Gliman, Alfred G.
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLIL CYCLASE
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,214
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/005,498
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:450
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3357 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-726-214-7

Query Match 2.6%; Score 77.8; DB 3; Length 3357;
Best Local Similarity 48.5%; Pred. No. 5e-11;
Matches 214; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

QY 1938 TGAAGCCAGAGTTGAGTAAATGTCACCATGCTCTCTCAGACATCGTTGGGTTCACTG 1997
DB 900 TGTATGTCAGAGGACCAAGGAGTGTGCTGTATGCTGACATCGTGGGCTTCACAC 959

QY 1998 CCATCTGCTCCCGTGTCAACCGTGCAGGTGCATCACCATGCTCAATGCTGACACTC 2057
DB 960 GGCTGGCCAGTGTGTTCCCTAAGGAGCTGGTCTAATGCTCAATGACTCTTCGGCA 1019

QY 2058 GCTTCGACAGAGTGTGGAGAGTGTCTACAGAGTGTGGAGACCATGCGGATGCCCT 2117
DB 1020 AATTCGACCAAAATTCGAAAGGAGCAGCAATGCAATGCGGATCAAGATCCTGGGAGCTGTT 1079

QY 2118 ATTGTGTAGTGGGGGATTAACAAAGAGAGTGTACTATGCTGTTCAGATAGCGGTGA 2177
DB 1080 ACTACTGTGTATCGGGCTGCCCTCTCTCTGCTGACACAGCCATCAATGCGTGGCA 1139

QY 2178 TGGCCCTTGAAGATGAGAGCTCTCTGATGAAGTTATGCTCCCCATGGAGAACCTATCA 2237
DB 1140 TGGAGCTGGACATGTGCGGGGCCATCAGGAACTTCGGGTAGCCACCGGTGTGGATATCA 1199

QY 2238 AGATGGGAATGAGTGCACCTCTGGATCAGTCTTTTGGTGGCGTCTGGAGTTAAATGC 2297
DB 1200 ACATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259

QY 2298 CCGGTACTCTCTTTTGGAAACAAATGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2357
DB 1260 GGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319

QY 2358 TACCACGAAATCAATGTCA 2378
DB 1320 TTCAGGACGAGTGCACATCA 1340

RESULT 13
US-08-726-214-17
; Sequence 17, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,214
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/005,498
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:450
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1652 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-726-214-17

Query Match 2.5%; Score 75.8; DB 3; Length 1652;
Best Local Similarity 48.5%; Pred. No. 1.2e-10;
Matches 209; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

QY 1958 AATGTCACCATGCTCTCTCAGACATCGTTGGGTTCACTGCCATCTGCTCCCAAGTGTCA 2017
DB 205 AACGTGAGCATCTCTTTGACAGACATCGTGGGCTTACAGGCTTGGCGTCACAGTGCACG 264

QY 2018 CCCTCCAGTGCATCACCATGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2077
DB 265 GCCCAGGAGTGTGTGAATCTCTCAATGAGCTCTTGGGAAGTTTACAGAGCTGGCCACA 324

QY 2078 GAGCTGATGCTACAGAGTGGAGACCATTTGGGATGCCCTATTTGTGTAGCTGGGGATTA 2137
DB 325 GAGAACCATGCTGCGCGCATCAGATCCTGGGAGATTGCTACTACTCGGTGTCTGGCCTC 384

QY 2138 CACAAAGAGAGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2197
DB 385 ACTCAGCCCAAGACTGACACGCGCCACTGCTGTGTGGAGATGGCGCTGGACATGATGCAC 444

QY 2198 CTCTCTGATGAAGTTATGCTCCCATGGAGACCTATCAAGATGCAATTTGCAATGCTGCT 2257
DB 445 ACCATCAGCTCCGCTGGCTGAGGCCACTGAGGTGAGCTTGAACATGCGTGGGGCTGCAC 504

QY 2258 TCTGGATCAGTTTTTGTGGCTGCTTGGAGTTAAATGCCCTTACTGCTCTTTTGTGA 2317
DB 505 ACCGCGAGGTTCTCTGCGGGTCTCTGCGGGTCCGCTAAGTGGCAGTATGATGTGTGGTCC 564

QY 2318 ACAAATGTCATCTGGCTACAAATTTGAGTCTCTGAGTGTACACGAGAAATCAATGTC 2377
DB 565 AACGAGTGTACCTGCGCAACGTCATGGAGGCTGCCGCGCTGCTGCGGAGGTTTCATC 624

QY 2378 AGCCCAACAC 2388
DB 625 ACAAGACCAC 635

RESULT 14
US-08-726-214-1
; Sequence 1, Application US/08726214

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 16:02:43 ; Search time 286.696 Seconds
(without alignments)
15611.019 Million cell updates/sec

Title: US-09-762-767A-1

Perfect score: 3015
Sequence: 1 cccttatggcgattggcggg.....aaaaaaaaaaaaaaaaaaaaa 3015

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2964	98.3	3017	9	US-10-205-823-161
2	2540	84.2	2850	9	US-10-205-823-163
3	2540	84.2	2871	9	US-10-205-823-164
4	1552	51.5	2430	9	US-09-952-213D-1
c 5	714.6	23.7	177556	9	US-09-952-213D-6
6	483.6	16.0	7697	9	US-09-952-213D-4
7	435	14.4	1037	9	US-10-198-846-13184
8	410.6	13.6	834	9	US-10-198-846-9332
9	308.4	10.2	852	9	US-10-198-846-6205
c 10	176.4	5.9	2335	10	US-09-917-800A-1569
11	127	4.2	7697	9	US-09-952-213D-4
12	121.4	4.0	3745	10	US-10-157-031-17
13	121.4	4.0	3787	10	US-09-819-249-1
14	91.4	3.0	4523	9	US-10-175-158-1
15	90.8	3.0	3518	9	US-10-121-911-2
16	77.6	2.6	3582	10	US-09-750-240-12
17	75.6	2.5	1812	10	US-09-750-240-3
18	75.6	2.5	3549	10	US-09-750-240-5
19	75.6	2.5	3552	10	US-09-750-240-10

Sequence 1, Appli	20	75.6	2.5	4942	9	US-10-201-000-1
Sequence 352, App	21	72	2.4	2601	10	US-09-925-297-352
Sequence 1, Appli	22	70.6	2.3	6196	9	US-10-282-942-1
Sequence 5, Appli	23	67	2.2	15093	9	US-09-952-213D-5
Sequence 6, Appli	24	65.4	2.2	177556	9	US-09-952-213D-6
Sequence 1, Appli	25	62.4	2.1	4985	12	US-10-071-223-1
Sequence 182, App	26	61.6	2.0	4473	10	US-09-751-100B-1
Sequence 98, Appl	27	60.8	2.0	5515	10	US-09-751-100B-98
Sequence 71, Appl	28	59.8	2.0	330	9	US-09-764-868-182
Sequence 445, App	29	59.8	2.0	330	9	US-09-989-442-71
Sequence 445, App	30	59.8	2.0	330	9	US-10-091-504-445
Sequence 445, App	31	59.8	2.0	330	10	US-09-764-869-445
Sequence 6274, Ap	c 32	55.2	1.8	418	9	US-09-796-692-6274
Sequence 6274, Ap	c 33	55.2	1.8	418	9	US-10-040-862-6274
Sequence 1351, Ap	34	51.4	1.7	11881	9	US-09-764-868-1351
Sequence 1353, Ap	35	51.4	1.7	11881	9	US-09-764-868-1353
Sequence 2124, Ap	36	51.4	1.7	11881	9	US-10-091-504-2124
Sequence 2124, Ap	37	51.4	1.7	11881	10	US-09-764-869-2124
Sequence 32338, A	38	49.2	1.6	459	9	US-09-918-995-32338
Sequence 30, Appl	39	49	1.6	778	9	US-09-774-639-30
Sequence 38, Appl	40	49	1.6	778	9	US-09-969-730-38
Sequence 95, Appl	41	48.8	1.6	1073	9	US-10-063-547-95
Sequence 305, App	42	48.8	1.6	1073	9	US-10-174-590-305
Sequence 305, App	43	48.8	1.6	1073	9	US-10-176-758-305
Sequence 95, Appl	44	48.8	1.6	1073	9	US-10-063-616-95
Sequence 305, App	45	48.8	1.6	1073	9	US-10-175-737-305

ALIGNMENTS

RESULT 1

US-10-205-823-161
; Sequence 161, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 3017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-161

Query Match 98.3%; Score 2964; DB 9; Length 3017;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 3011; Conservative 0; Mismatches 0; Indels 7; Gaps 4;

QY	4	TTATGGCGATTGGCGGCTCGAGAGACCAGGACTCAGTTCCCTTGCCTTACTCTGAGCCT	63
DB	1	TTATGGCGATTGGCGGCTCGAGAGACCAGGACTCAGTTCCCTTGCCTTACTCTGAGCCT	60
QY	64	AGTGGGTGGGACTCAGCTCAGAGTCAGTTTTTCAGAAAGCAGGTTTTCAGTGCAGAGTTTTC	123
DB	61	AGTGGGTGGGACTCAGCTCAGAGTCAGTTTTTCAGAAAGCAGGTTTTCAGTGCAGAGTTTTC	120
QY	124	CTACACTTTTCTCGCTAGAGCAGCAGCAGCCTTGGAAACAGACCCAGGGGGAGGACACC	183
DB	121	CTACACTTTTCTCGCTAGAGCAGCAGCAGCCTTGGAAACAGACCCAGGGGGAGGACACC	180
QY	184	TGTGGGGGAGCGGCTCGAGAGCTTAGAGACCCAGCCCGGGCGTGTATCTCAACCATG	243
DB	181	TGTGGGGGAGCGGCTCGAGAGCTTAGAGACCCCGGGCGTGTATCTCAACCAT -	239
QY	244	TGGGGATTTCGAGGGCGGCTTGGAGTCTTAGAGATCCGGAAGCACAGCCCCGAGGTG	303
DB	240	TGGGGATTTCGAGGGCGGCTTGGAGTCTTAGAGATCCGGAAGCACAGCCCCGAGGTG	299
QY	304	TGGGAAGCACCAAGACTTGGGCTCTTGGAGAAAGCGTAGAGGGGGCCACCGCGGTCT	363
DB	300	TGGGAAGCACCAAGACTTGGGCTCTTGGAGAAAGCGTAGAGGGGGCCACCGCGGTCT	359
QY	364	CCGGCTGTCTGCACCTCTCGCCTGAGCTGCTAGAGATCCGGAAGCACAGCCCCGAGGTG	423
DB	360	CCGGCTGTCTGCACCTCTCGCCTGAGCTGCTAGAGATCCGGAAGCACAGCCCCGAGGTG	419
QY	424	CCAGTGTCTTGAATTGATAGTGGCTTCTGTTGTAGTCTCATATAGAAGTACAGCTC	483
DB	420	CCAGTGTCTTGAATTGATAGTGGCTTCTGTTGTAGTCTCATATAGAAGTACAGCTC	479
QY	484	ATCAGGAGGAGATCGCAGCAGGCTAGAGACACCAACACCATGTTCTGCAGAGCTCAA	543
DB	480	ATCAGGAGGAGATCGCAGCAGGCTAGAGACACCAACACCATGTTCTGCAGAGCTCAA	539
QY	544	GGATCTCAAGTACAGGAGAGTGCTCTTCTTACTGACACAGGCTCAAGTTCCTAA	603
DB	540	GGATCTCAAGTACAGGAGAGTGCTCTTCTTACTGACACAGGCTCAAGTTCCTAA	599
QY	604	CGAGTCTTCAGAGGAGCAGCAGGAGCTCAGAGAGCTGCAAGACCAACCGTGCCCATCTG	663
DB	600	CGAGTCTTCAGAGGAGCAGCAGGAGCTCAGAGAGCTGCAAGACCAACCGTGCCCATCTG	659
QY	664	TCAAGACATTCCTGAGAGAAACATACAGAAAGTCTTCCTCAAGAAAAACAGTCGGAG	723
DB	660	TCAAGACATTCCTGAGAGAAACATACAGAAAGTCTTCCTCAAGAAAAACAGTCGGAG	719
QY	724	CCGAGTCTATCTTTCACACTTTGACAGAGAGTATTTGCAAACTGATTTTCCACAGTTTGA	783
DB	720	CCGAGTCTATCTTTCACACTTTGACAGAGAGTATTTGCAAACTGATTTTCCACAGTTTGA	779
QY	784	ACGGCTGAATGTTGCACTTCAGAGAACATTTGGCAAGACAAAAATAAGAAAGACAGGAA	843
DB	780	ACGGCTGAATGTTGCACTTCAGAGAACATTTGGCAAGACAAAAATAAGAAAGACAGGAA	839
QY	844	ATCTTTTGGAAAGAGAGACTTTGAAAAAAACAATTCAGAGCAGAGCTTCGACGAGAGT	903
DB	840	ATCTTTTGGAAAGAGAGACTTTGAAAAAAACAATTCAGAGCAGAGCTTCGACGAGAGT	899
QY	904	TCCAGTGGAGGTTTATCAAGAAATCTCTTGGTGAAGAGGTTTTTAAATATGTTTACAGGA	963
DB	900	TCCAGTGGAGGTTTATCAAGAAATCTCTTGGTGAAGAGGTTTTTAAATATGTTTACAGGA	959
QY	964	AGATGAAACATCCTTTGGGCTGTTGGAGCACCCCTTAAAGATTTTTTAAACAGCTTCAG	1023
DB	960	AGATGAAACATCCTTTGGGCTGTTGGAGCACCCCTTAAAGATTTTTTAAACAGCTTCAG	1019
QY	1024	TACCTTCTGAAACAGAGCAGCATTTGCCAAGAACAGGAAAAAGGGCAGGCTTGAGGA	1083
DB	1020	TACCTTCTGAAACAGAGCAGCATTTGCCAAGAACAGGAAAAAGGGCAGGCTTGAGGA	1079

```
QY 2164 TCAGATAGCGCTGATGGCCCTGAAGATGATGAGGCTCTCTGATGAAGTATATCTCCCA 2223
Db 2160 TCAGATAGCGCTGATGGCCCTGAAGATGATGAGGCTCTCTGATGAAGTATATCTCCCA 2219
QY 2224 TGGAGAACCTATCAAGATCGGAATGGAGCTGACCTCTGGATCAGTCTTTGCTGGGCTCT 2283
Db 2220 TGGAGAACCTATCAAGATCGGAATGGAGCTGACCTCTGGATCAGTCTTTGCTGGGCTCT 2279
QY 2284 TGGAGTAAATGCCCGTACTGCTCTTTTGGAAACAAATGTCACCTCTGGCTTAACAAAT 2343
Db 2280 TGGAGTAAATGCCCGTACTGCTCTTTTGGAAACAAATGTCACCTCTGGCTTAACAAAT 2339
QY 2344 TCAGTCTCGAGCTACACGAGAAATCAATGTACGCCCAACAACTACAGATTAACAA 2403
Db 2340 TCAGTCTCGAGCTACACGAGAAATCAATGTACGCCCAACAACTACAGATTAACAA 2399
QY 2404 AGACTGTCTGCTGCTGCTGCTTTACCCCTGCAATCAAGGAGGAACTTCCACCAACTTCC 2463
Db 2400 AGACTGTCTGCTGCTGCTTTACCCCTGCAATCAAGGAGGAACTTCCACCAACTTCC 2459
QY 2464 TAGTGAATCCCGGAATGTCGCAATTTCTGGATGCTTACCAACAAAGGAACAACTCAA 2523
Db 2460 TAGTGAATCCCGGAATGTCGCAATTTCTGGATGCTTACCAACAAAGGAACAACTCAA 2519
QY 2524 ACCATGCTTCCAAAGAAAGATGTGAAGATGGCAATGCAATTTTATAGGCAAGCATC 2583
Db 2520 ACCATGCTTCCAAAGAAAGATGTGAAGATGGCAATGCAATTTTATAGGCAAGCATC 2579
QY 2584 AGAATAGATTAGCAACCTATATACCTATTTAAGTCTTTGGGGTTTACATTTGAAG 2643
Db 2580 AGAATAGATTAGCAACCTATATACCTATTTAAGTCTTTGGGGTTTACATTTGAAG 2639
QY 2644 ATGTGTAGAGCTCTGAAAGCACTTTAGGATTTGTAGATGGCTTAACAAAGCACTATTA 2703
Db 2640 ATGTGTAGAGCTCTGAAAGCACTTTAGGATTTGTAGATGGCTTAACAAAGCACTATTA 2699
QY 2704 TTTTACGAGCCAGTCACATCTTCTCTGTTTACATGACAA- - - - -ATGTACTACT 2759
Db 2700 TTTTACGAGCCAGTCACATCTTCTCTGTTTACATGACAAATGATGTACTACT 2759
QY 2760 TCAGTACTTTCAGCTCTTCAAG- - - - -AAAAAAGCTTTTAAAGCTTTTGGGAG 2818
Db 2760 TCAGTACTTTCAGCTCTTCAAGAAAAAAGCTTTTAAAGCTTTTGGGAG 2819
QY 2819 TATTTCTATATATACAGCACTTACTACCTGTACTCAAAATACAGCACTTTGTACATA 2878
Db 2820 TATTTCTATATATACAGCACTTACTACCTGTACTCAAAATACAGCACTTTGTACATA 2879
QY 2879 TATCAGATATTTAGTCAATTTGTACAAACTGATGGAGTCACTGCAATCTCATATCTTG 2938
Db 2880 TATCAGATATTTAGTCAATTTGTACAAACTGATGGAGTCACTGCAATCTCATATCTTG 2939
QY 2939 GTGGATGCCATGTTTAAAGTGTGTTTGTGATAG- - - - -TGCTGCAAAAAA 2997
Db 2940 GTGGATGCCATGTTTAAAGTGTGTTTGTGATAGTGTGTTGATAGTGTGCTCA 2999
QY 2998 AAAAAAAAAAAAAAAAAA 3015
Db 3000 AAAAAAAAAAAAAAAAAA 3017
```

RESULT 2

```
US-10-205-823-163
; Sequence 163, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
```

```
; APPLICANT: Wensey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163
; LENGTH: 2850
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-205-823-163
```

```
Query Match 84.2%; Score 2540; DB 9; Length 2850;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2565; Conservative 0; Mismatches 0; Indels 5; Gaps 2;

QY 412 ACATCCAGTACCAGTGTCTGAATGATAGTGGCTTCTGTTTGTCTCATATATAA 471
Db 274 ACATCCAGTACCAGTGTCTGAATGATAGTGGCTTCTGTTTGTCTCATATAA 333
QY 472 GAATACAGTCTATCAGGAGGAGATCGCAGCAGGGAAGAGACACCAACATGTCTG 531
Db 334 GAATACAGTCTATCAGGAGGAGATCGCAGCAGGGAAGAGACACCAACATGTCTG 393
QY 532 CACGAAGCTCAAGGATCTCAAGATCACAGGAGTGTCTTCTCTCTTACTGGCACCAG 591
Db 394 CACGAAGCTCAAGGATCTCAAGATCACAGGAGTGTCTTCTCTCTTACTGGCACCAG 453
QY 592 TCAAGTTCCTAACAGTCTTTCAGAGGAGCAGCAGGAGCTCAGAGAGCTGCAAGCAAC 651
Db 454 TCAAGTTCCTAACAGTCTTTCAGAGGAGCAGCAGGAGCTCAGAGAGCTGCAAGCAAC 513
QY 652 CGTGCCCATCTGTCAAGACATTCCTCGAAGAACATACAAAGATCTTCTCTCAAGAA 711
Db 514 CGTGCCCATCTGTCAAGACATTCCTCGAAGAACATACAAAGATCTTCTCTCAAGAA 573
QY 712 AACAGTTCGAGCGAGTCTATCTTCCACTTTGGCAGAGATTTTGAACACTGATTTT 771
Db 574 AACAGTTCGAGCGAGTCTATCTTCCACTTTGGCAGAGATTTTGAACACTGATTTT 633
QY 772 CCCAGAGTTTGAACGGCTGAATGTTGCATTTCAAGAACATTTGGCAAGCACAATAA 831
Db 634 CCCAGAGTTTGAACGGCTGAATGTTGCATTTCAAGAACATTTGGCAAGCACAATAA 693
QY 832 AGAAGCAGGAATCTTTTGAAGAGAGAGCTTTGAAAACAACTTTCAGAGCAGCAGT 891
Db 694 AGAAGCAGGAATCTTTTGAAGAGAGAGCTTTGAAAACAACTTTCAGAGCAGCAGT 753
QY 892 TGCAGCAGAGTCCAGTGGAGTTATCAAGATCTTCTTGGTGAAGAGGTTTTTAAAT 951
Db 754 TGCAGCAGAGTCCAGTGGAGTTATCAAGATCTTCTTGGTGAAGAGGTTTTTAAAT 813
QY 952 ATGTTACGAGGAAGATGAAAACATCTTGGGGTGTGGAGGACCCCTTAAAGATTTT 1011
Db 814 ATGTTACGAGGAAGATGAAAACATCTTGGGGTGTGGAGGACCCCTTAAAGATTTT 873
QY 1012 AAACAGCTTCAGTACCCTTCTTGAACAGAGCAGCCATTTGCCAAGAGAGGAGG 1071
```

Db 874 ANACAGCTTTCAGTACCTTCTGAACACAGAGCAGCCATTGGCCAAAGACGAGAAAAGGGG 933
QY 1072 CAGGCTTGAGGACGCTCCATCTTATGCTGGATGAAGAGAGATGATTTTCTACATGTTTA 1131
Db 934 CAGGCTTGAGGACGCTCCATCTTATGCTGGATGAAGAGAGATGATTTTCTACATGTTTA 993
QY 1132 CTACTTCTTCCCTAAGAGAACCACTCCCTGATTTCTCCGGGCATCAATAAGGACGCTGC 1191
Db 994 CTACTTCTTCCCTAAGAGAACCACTCCCTGATTTCTCCGGGCATCAATAAGGACGCTGC 1053
QY 1192 TCAGTATTATATGAACGGAAGTGAAGTGTCTTAATGCTCCCTGCTTCATATGA 1251
Db 1054 TCAGTATTATGAACGGAAGTGAAGTGTCTTAATGCTCCCTGCTTCATATGA 1113
QY 1252 TTGAGGAGAGTTTGTGAATCAGCCCTACTTGTGTCTACTCCGTTCCACATGAAGACACCAA 1311
Db 1114 TTGAGGAGAGTTTGTGAATCAGCCCTACTTGTGTCTACTCCGTTCCACATGAAGACACCAA 1173
QY 1312 GCCATCCCTGTCCCGACGAAACCCCGATCTCGCTGGTGNATCCCACTGCTATTCTG 1371
Db 1174 GCCATCCCTGTCCCGACGAAACCCCGATCTCGCTGGTGNATCCCACTGCTATTCTG 1233
QY 1372 CAAGACATTTCCATTTCCATTTTCATGTTTGAACAGATATGCAATTTCTGCAATTTGCAA 1431
Db 1234 CAAGACATTTCCATTTCCATTTTCATGTTTGAACAGATATGCAATTTCTGCAATTTGCAA 1293
QY 1432 TGGCATCAGAGGCTGATGAACAGGAGAGACTTTCAAGGAAGGCTTAATTTGAAGAATA 1491
Db 1294 TGGCATCAGAGGCTGATGAACAGGAGAGACTTTCAAGGAAGGCTTAATTTGAAGAATA 1353
QY 1492 CTTTGAATTTCTGACTCAAAATCAACAGAGAGCTTTAGCGGATCATGACTATGTTGAA 1551
Db 1354 CTTTGAATTTCTGACTCAAAATCAACAGAGAGCTTTAGCGGATCATGACTATGTTGAA 1413
QY 1552 TATCAGTTTGTGTACAGTGAAGAGATGGCAACTCTGTGAAGAACTTCAAGGCT 1611
Db 1414 TATCAGTTTGTGTACAGTGAAGAGATGGCAACTCTGTGAAGAACTTCAAGGCT 1473
QY 1612 TATGAGCTCAAGGCCAAATGATCTACATTTGTTGAATCCAGTCAATCTGTTTGGG 1671
Db 1474 TATGAGCTCAAGGCCAAATGATCTACATTTGTTGAATCCAGTCAATCTGTTTGGG 1533
QY 1672 GTCACCTGTGTGACAGATAGAGATTTTACAGGACGAGGCTTACCTCTCAGACAT 1731
Db 1534 GTCACCTGTGTGACAGATAGAGATTTTACAGGACGAGGCTTACCTCTCAGACAT 1593
QY 1732 CCCAATTCACAATCAGTCAAGGATGTGGTCTTAATAGGGGAACAAGCCGAGCTCAAGA 1791
Db 1594 CCCAATTCACAATCAGTCAAGGATGTGGTCTTAATAGGGGAACAAGCCGAGCTCAAGA 1653
QY 1792 TGGCCTGAAGAGAGGCTGGGAAGCTGAAGGCTACCCCTTGAGCAAGCCCAAGCCCT 1851
Db 1654 TGGCCTGAAGAGAGGCTGGGAAGCTGAAGGCTACCCCTTGAGCAAGCCCAAGCCCT 1713
QY 1852 GGAGGAGAGAGAGAGAGACAGTACCTCTGTGCTCCATATTTCCCTGTGAGGTGC 1911
Db 1714 GGAGGAGAGAGAGAGAGACAGTACCTCTGTGCTCCATATTTCCCTGTGAGGTGC 1773
QY 1912 TCAGCAGCTGTGCAAGGCAAGTGTGTGAAGCAAGCAAGTTCAGTAAATGTCAACCATGCT 1971
Db 1774 TCAGCAGCTGTGCAAGGCAAGTGTGTGAAGCAAGCAAGTTCAGTAAATGTCAACCATGCT 1833
QY 1972 CTTCTCAGACATCGTTGGGTTCACTGCCATCTGCTCCAGTGTCTACCCGTGAGGTCAT 2031
Db 1834 CTTCTCAGACATCGTTGGGTTCACTGCCATCTGCTCCAGTGTCTACCCGTGAGGTCAT 1893
QY 2032 CACCATGCTCAATGCACTGTACACTCCCTCGACAGCAGTGTGGAGAGCTGATGTCTA 2091
Db 1894 CACCATGCTCAATGCACTGTACACTCCCTCGACAGCAGTGTGGAGAGCTGATGTCTA 1953
QY 2092 CAAGGTGGAGACCAATGGCGATGCTATTGTAGTGGGGATTACACAAAGAGAGTGA 2151

Db 1954 CAAGGTGGAGACCAATGGCGATGGCTATTTGTAGTGGGGATTACACAAAGAGTGA 2013
QY 2152 TACTCATGCTGTTTCAGATAGCGCTGATGGCCCTGAAGATGATGAGGCTCTCTGATGAAGT 2211
Db 2014 TACTCATGCTGTTTCAGATAGCGCTGATGGCCCTGAAGATGATGAGGCTCTCTGATGAAGT 2073
QY 2212 TATGCTCTCCCATGGAGAACCTATCAAGATGCGAATTTGGACTGCACTCTGGATCAGTTTT 2271
Db 2074 TATGCTCTCCCATGGAGAACCTATCAAGATGCGAATTTGGACTGCACTCTGGATCAGTTTT 2133
QY 2272 TGCTGGCGTGTGGAGCTTAAATGCCCCCTTACTGTCTTTTGGAAACAAATGTCACCTCT 2331
Db 2134 TGCTGGCGTGTGGAGCTTAAATGCCCCCTTACTGTCTTTTGGAAACAAATGTCACCTCT 2193
QY 2332 GGCTAAACAAATTTGAGTCTCGAGTGTACACAGAAAATCAATGTACGCCCAACAACTTA 2391
Db 2194 GGCTAAACAAATTTGAGTCTCGAGTGTACACAGAAAATCAATGTACGCCCAACAACTTA 2253
QY 2392 CAGATTAATAAGACTGTCTGCTGTTTGTGTTTACCCCTCGATCAAGGAGGAACTTCC 2451
Db 2254 CAGATTAATAAGACTGTCTGCTGTTTGTGTTTACCCCTCGATCAAGGAGGAACTTCC 2313
QY 2452 ACCAAACTCTCCCTAGTGAATCCCGGAATCTGCCATTTCTGGATGCTTACCACCAAGG 2511
Db 2314 ACCAAACTCTCCCTAGTGAATCCCGGAATCTGCCATTTCTGGATGCTTACCACCAAGG 2373
QY 2512 AACAACTCAAAACCATGCTTCCAAAAGAAAGATGTGAAGATGGCAATGCCAATTTTTT 2571
Db 2374 AACAACTCAAAACCATGCTTCCAAAAGAAAGATGTGAAGATGGCAATGCCAATTTTTT 2433
QY 2572 AGGCAAGCATCAGGAATAGATTAGCAACCTATATACCTATTTATAGTCTTTGGGGTTT 2631
Db 2434 AGGCAAGCATCAGGAATAGATTAGCAACCTATATACCTATTTATAGTCTTTGGGGTTT 2493
QY 2632 GACTCATTTGAGATGCTGAGGCTCTGAAAGCACTTTAGGGATGTAGATGGCTACAA 2691
Db 2494 GACTCATTTGAGATGCTGAGGCTCTGAAAGCACTTTAGGGATGTAGATGGCTACAA 2553
QY 2692 GCAGTATTAATAATTTCAAGGACCAAGTCACAATCTTTCTCTCTTTTAAACATGACAAA --- 2748
Db 2554 GCAGTATTAATAATTTCAAGGACCAAGTCACAATCTTTCTCTCTTTTAAACATGACAAA 2613
QY 2749 -ATGCTACTACTTCAGTACTTTCAGCTCTTCAAG-AAAAAANAACCTTTAAAAAGCTA 2806
Db 2614 TATGCTACTACTTCAGTACTTTCAGCTCTTCAAGAAAAAANAACCTTTAAAAAGCTA 2673
QY 2807 CTTTGTGGGAGTATTTCTATATATAACAGCACTTACTACCTGTACTCAAAATTCAGC 2866
Db 2674 CTTTGTGGGAGTATTTCTATATATAACAGCACTTACTACCTGTACTCAAAATTCAGC 2733
QY 2867 ACCTGTGACATATATCAGATTAATTTAGTCAATTTGTAACTGATGGAGTACCTGCAA 2926
Db 2734 ACCTGTGACATATATCAGATTAATTTAGTCAATTTGTAACTGATGGAGTACCTGCAA 2793
QY 2927 TCTCATATCTGCTGGATGCCATGTTTAAAGTGTGTTGTGATAGT 2976
Db 2794 TCTCATATCTGCTGGATGCCATGTTTAAAGTGTGTTGTGATAGT 2843

RESULT 3

US-10-205-823-164

; Sequence 164, Application US/10205823

; Publication No. US20030108963A1

; GENERAL INFORMATION:

; APPLICANT: Schlegel, Robert

; APPLICANT: Monahan, John E.

; APPLICANT: Endege, Wilson O.

; APPLICANT: Gannavarapu, Manjula

; APPLICANT: Gorbacheva, Bella

; APPLICANT: Hoersch, Sebastian

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Womsey, Angela M.

; APPLICANT: Glatt, Karen

Db 24584 CCCTGTCCCCAGCAACCCAGTCTCGTGGTGAATCCACATCGCTATTCTGCAAGA 24525
QY 1377 CATTTCATTCCTTCATCTTTGACAAACATATGACAATTCGCAATTTGGCAATGGCA 1436
Db 24524 CATTTCATTCCTTCATCTTTGACAAACATATGACAATTCGCAATTTGGCAATGGCA 24465
QY 1437 TCAGAAGGCTGTGAACAGGAGAGACTTTTCAAGGAAGCCCTAAATTTTGAAGAATACTTTG 1496
Db 24464 TCAGAAGGCTGTGAACAGGAGAGACTTTTCAAGGAAGCCCTAAATTTTGAAGAATACTTTG 24405
QY 1497 AATTCCTGACTCCAAAATCAACAGAGCTTTTACGGGATCATGATCTTTGAATATGC 1556
Db 24404 AATTCCTGACTCCAAAATCAACAGAGCTTTTACGGGATCATGATCTTTGAATATGC 24345
QY 1557 AGTTTGTGTACGAGTGAGGAGATGGGACAACCTCTGTGAAGAAATCTTCAAGGGTATGG 1616
Db 24344 AGTTTGTGTACGAGTGAGGAGATGGGACAACCTCTGTGAAGAAATCTTCAAGGGTATGG 24285
QY 1617 A 1617
Db 24284 A 24284

RESULT 6

US-09-952-213D-4
; Sequence 4, Application US/09952213D
; Publication No. US20030096240A1
; GENERAL INFORMATION:
; APPLICANT: MURAD, FERID
; APPLICANT: SHARINA, IRAIDA G.
; APPLICANT: KRUMENACKER, J. S.
; APPLICANT: MARTIN, E.
; TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC
; FILE REFERENCE: UFSH:252US
; CURRENT APPLICATION NUMBER: US/09/952,213D
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 7697
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (605)..(6955)
; OTHER INFORMATION: N = A, C, T/U OR G
US-09-952-213D-4

Query Match 16.0%; Score 483.6; DB 9; Length 7697;
Best Local Similarity 78.9%; Pred. No. 7.5e-127;
Matches 584; Conservative 3; Mismatches 152; Indels 1; Gaps 1;
QY 896 GCAGGAGTTCAGTGAGGTTATCAAGAATCTCTGGTGAAGAGGTTTTTAAATATGT 955
Db 2862 GCAGGTGCCCGTGGAGGCGCTCAAGACTCTCTGGCGGAGGAGCTGTTCAAGATCTGC 2921
QY 956 TAGCAGGAAGATGAACATCTCTGGGTGGTGGAGGACCTTAAAGATTTTTTAAAC 1015
Db 2922 TAGCAGGAAGATGACACATTTTGGCGTGGTGGCGGACCTTGAAGGACTTCTTAAT 2981
QY 1016 AGTTTCAGTACCCTTCTGAAACAGAGCAGCCATTTGCCAAGACAGGAAAGGGCAGG 1075
Db 2982 AGTTTCAGCAGCTCTCTCAAGCAGAGCAGCAGCTGCCAAGAGGCGGAGGCGGACGA 3041
QY 1076 CTTGAGGACCCCTCCATCTCTATGCTGGTAAGAGGATGATTTTCTACATGTTTACTAC 1135
Db 3042 CTTGAGATGCCCTCCATCTTATGCTGGACAGGACAGGACTTCTTAAATGTTTACTAC 3101
QY 1136 TTCTTCCCTAAGAGAACCACTCCCTGATTTCTCCCGGCATCATAAAGGCGAGCTGCTCAC 1195
Db 3102 TTCTTCCCGAAGAGAACCAAGCCCTGCTTCTCTCCCTGGTATCATTAAGCGGCTGCTCGC 3161

QY 1196 GTATTATATGAACGGAAGTGAAGTGTCTTAAATGCCCTCCCTGCTTCCATAATGATTCG 1255
Db 3162 ATACTGTAGAAAGCCAGCTGGAGGTGCTCCTGATGCTCCTGCTTCCGAAGTGAAGTGT 3221
QY 1256 AGCGAGTTTGTGAATCAGCCCTACTTGTGTACTCCGTTTCACATGAAACGACCAAGCCA 1315
Db 3222 ACCGAGTTTGTGAACGAGCCCTATTGCTCTACTCCGTTTCATGTAAGAGCAGCAAGCCS 3281
QY 1316 TCCTCTGCCCCAGCAACCCAGTCTCTCGTGGTATTCACACATGCTGCTTATCTGC-AA 1374
Db 3282 TCCTCTGCCCCAGCAACCCAGTCTCTCGTGGTATTCACACATGCTGCTTATCTGC-AA 3341
QY 1375 GACATTTCCATTTCCATTTTTCATGTTTGAACAAAGATATGACAATTTGCAATTTGGCAATGG 1434
Db 3342 GACTTTCCCGTTTSCATTTTCATGCTGGACCGAGACCTGGCCATCTGCGAGCTGGTACGG 3401
QY 1435 CATCAGAAGGCTGATGAACAGGAGAGACTTTTCAAGGAAGCCCTAAATTTTGAAGAACTTT 1494
Db 3402 CATCAGAAGGCTGCTGAACAAAGAGGAGCTTCCAAAGGGAAGCCCACTTTTGAAGAGTTCTT 3461
QY 1495 TGAATTTCTGACTCCCAAAATCAACAGAGCTTTTAGCGGGATCATGACTATGTTGAATAT 1554
Db 3462 TGAATTTCTAATCTCCCAAAATCAACAGAGACTTTTAGGGATCATGACATGTTGATAT 3521
QY 1555 GCAGTTTGTGTACGAGTGAAGAGATGGGAAACTCTGTGAAGAAATCTTCAAGGGTAT 1614
Db 3522 GCAGTTTGTTCATCCGGGTGAGGAGATGGGATACTCGGTGAAGAAATCGTCAAGGGTAA 3581
QY 1615 GGACCTCAAGGCAATGA 1634
Db 3582 GAAATGTAAACGGGATTC 3601

RESULT 7

US-10-198-846-13184
; Sequence 13184, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13184
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1037
; OTHER INFORMATION: n = A, T, C or G
US-10-198-846-13184

Query Match 14.4%; Score 435; DB 9; Length 1037;
Best Local Similarity 98.0%; Pred. No. 1.5e-113;
Matches 483; Conservative 0; Mismatches 5; Indels 5; Gaps 4;
QY 412 ACATCCAGTTACCACTGCTTGAATGATGAGTGGCTTCTGTTTGTGAGTCTCATATAA 471
Db 210 ACATCCAGTTACCACTGCTTGAATGATGAGTGGCTTCTGTTTGTGAGTCTCATATAA 269
QY 472 GAACACTACGCTATCAGGAGGATCGCAGCGGTAAGAGACCAACCAACCATGTTCTG 531
Db 270 GAACACTACGCTATCAGGAGGATCGCAGCGGTAAGAGACCAACCAACCATGTTCTG 329

649	QY	AACCGTCCCATCTGTCTCAAGACATPCTCTGAGAGAAACATACAGAAGAGTCTTCTCTCAAG	708
491	Db	AACCGTCCCATCTGTCTCAAGACATPCTCTGAGAGAAACATACAGAAGAGTCTTCTTAAAG	550
709	QY	AAAAACCAAGT-CGGAGCGGAGTCTATCTTCAACATTTGGCA-GAGAGTATTTTGCACAACTG	766
551	Db	AAAAACCAAGTCCGGAGCGGAGTCTATCTTACACTTTGGCANGAAAGTATTTTCAAACCTG	610
767	QY	ATTTTCCACAG-TTTGAACGGCTGAA-TCCTGCACCTCAGAA	809
611	Db	ATTTTCCACAGATTTTGAACGGCTGAATCTGTACACTTCANAAA	655

RESULT 10

```

US-09-917-800A-1569
; Sequence 1569, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.

```

; TITLE OF INVENTION: Molecular Toxicology Modeling

```

; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1569
; LENGTH: 2335
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012770
; US-09-917-800A-1569

```

Db	1201	TGTCCTGCCAACTGGAAAAGAAGAGGAGCTTGGTGTCTCTTTCCAATCACCTGGCCA	1260
Qy	1857	AGGAGAAGAAAAGACAGCTAGACCTTCTGTGCTCCATATTTCCCTGTGAGGTGCTCAGC	1916
Db	1261	TCGAGAAGAAGAAGACAGAGACCTTCCGTGTAGCCCATGCTGCCTGAACATGTGGCCAACC	1320
Qy	1917	AGCTGTGGCAAGGGCAAGTTGTGCAAGCCACAAGTTCAGTAAATGTCACATGCTCTTCT	1976
Db	1321	NACTCAAGGAGGGCAGAAAGGTGGCTGCAGAGAGAAATTTGAACATGTACAATCCTTTTCA	1380
Qy	1977	CAGACATCTGTGGTTTCACTGCATCTGTCTCCAGTGTCAACGCTGCAGGTCACTACCA	2036
Db	1381	CGCATGTTGTGACATTTACCAACATCTGTCGACGCTGTCAACCTATCCAATCTGGAACA	1440
Qy	2037	TGCTCAATGCATGTTACATCTCGTTGCGACACAGCTGTGGAGAGTGGATGTCCTACAAGG	2096
Db	1441	TGCTGTAATTCATGTACTCCAAGTTTGACAGGTTAAACCAAGTGTCCATGATGCTACAAAG	1500
Qy	2097	TGGAGACCATTGGCATGCTTATGTGTAGCTGGGGATTACACAAAGAGAGTGTACTCTC	2156
Db	1501	TAGAACAAATGGGGATGCTTACATGGTGGTGGGTGAGTCCAGTACCCGTTGAAAGCC	1560
Qy	2157	ATGCTGTTCCAGATAGCGCTGATGGCCCTGAAGATGATGGAGCTCTCTGATGAAGTATGT	2216
Db	1561	ATGCTCAAAAGAGTCGCCAAATTTTGCTCTGGGATGAGAATTTCTGCAAAAAGAGTGATGA	1620
Qy	2217	CTCC---CCATGGAGAACCTATCAAGATGCGAATTCGAGTGCACCTCTCGATCAGTTTTTG	2273
Db	1621	ATCCTGTCACTGGGGAACTATCCAGATCAGAGTGGGAATCCCACTGGACCACTGTAG	1680
Qy	2274	CTGGGCTCGTTGGAGTTAAATGCCCGTTACTGTCTTTTGGAAACAATGTCACTCTGG	2333
Db	1681	CAGGTGTTGTGGGAGACAAGATGCTCGGTACTGCTGTTGTTGGTGACACTGTAAACACAG	1740
Qy	2334	CTAACAAATTTGAGTCCCTGCATGTACACAGAAAATCAATCTCAGGCCCACTTACA	2393
Db	1741	CCTCTAGGATGGAAGTCAAGGGCTTCCAGCAAGTGCATCTGAGCCCCACAGCCCCACA	1800
Qy	2394	GATTACTCAAGAC	2407
Db	1801	GAGCCCTGAAAAAC	1814

RESULT 11

```

RESULT 11
US-09-952-213D-4/c
; Sequence 4, Application US/09952213D
; Publication No. US20030096240A1
; GENERAL INFORMATION:
; APPLICANT: MURAD, FERID
; APPLICANT: SHARINA, IRAIDA G.
; APPLICANT: KRUMENACKER, J. S.
; APPLICANT: MARTIN, E.
; TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC
; FILE REFERENCE: US/25205
; CURRENT APPLICATION NUMBER: US/09/952,213D
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 7697
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (605)..(6955)
; OTHER INFORMATION: N = A, C, T/U OR G
US-09-952-213D-4

```

Query: Match 4.28: Score 127: DB 9: Length 7697:

Query Match 4.26; Score 127; DB 9; Length 7697;
Best Local Similarity 87.48; Pred. No. 8.7e-25;

BEST LOCAL SIMILARITY 87.4%; Pred. NO: 8./E-23;
Matches 139: Conservative 0: Mismatches 20: Indels

Matches 139; Conservative 0; Mismatches 20; Indels

c	1	2443	100.0	2443	21	AAZ88939	Human soluble guan
2	2423	99.2	3196	22	AAH98392	Human EST-derived	
3	570	23.3	1302	22	AAH33895	Human colon cancer	
4	434.8	17.8	459	21	AAAC00760	Human secreted pro	
5	371.4	15.2	2850	23	ABL02665	Drosophila melanog	
6	200.2	8.2	2954	24	AA594827	Human DNA sequence	
7	192.6	7.9	2715	21	AAZ51684	Human cyclic nucle	
8	192.6	7.9	3015	23	AAZ88938	Human soluble guan	
9	191	7.8	4742	23	ABV21348	Human prostate exp	

PT New human soluble guanylate cyclase alpha/beta1 and the nucleic acid
 PT encoding the subunits, useful for producing diagnostic antibodies, and
 PT for somatic gene therapy of arteriosclerosis -
 XX
 PS
 PS
 XX Example 1; Page 15-16; 44pp; German.

CC This invention describes novel purified human soluble guanylate cyclase
 CC alpha/beta1 (hsGcalphal/beta1). The products of the invention have
 CC antihypertensive, vasotropic and hypotensive activity. Nucleic acid
 CC sequences encoding the alpha and/or beta subunit are useful for somatic
 CC gene therapy of arteriosclerosis and restenosis, ischemia (infarct),
 CC peripheral arterial occlusive disease and arterial hypertension.
 CC Antibodies to hsGcalphal/beta1 can be used for diagnosis of aberrant
 CC hsGcalphal/beta1 expression in human tissues. This sequence encodes the
 CC human soluble guanylylcyclase beta1 subunit described in the method
 CC of the invention.

XX
 SQ Sequence 2443 BP; 709 A; 513 C; 533 G; 688 T; 0 other;

Query Match 100.0%; Score 2443; DB 21; Length 2443;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCCCCCCCCGCGCTGCGCCCTGCTGGTGGTCCCTTGGCGGTACCTCTGGTGGGGC	60
DB	1	CCCCCCCCCGCGCTGCGCCCTGCTGGTGGTCCCTTGGCGGTACCTCTGGTGGGGC	60
QY	61	TGCTCCCGCGTCCCGGTGCAGACACCATGTACGGATTGTGAATCAGCCCTCGAGTT	120
DB	61	TGCTCCCGCGTCCCGGTGCAGACACCATGTACGGATTGTGAATCAGCCCTCGAGTT	120
QY	121	GCTGGTATCCGAATTAAGCGCCCGAGGTGGGAGACATCAAAAAGAGCAGATT	180
DB	121	GCTGGTATCCGAATTAAGCGCCCGAGGTGGGAGACATCAAAAAGAGCAGATT	180
QY	181	AGATGAAGAGCAGATTCTGTGAGATAATATATATATATATATATATATATATAT	240
DB	181	AGATGAAGAGCAGATTCTGTGAGATAATATATATATATATATATATATATATAT	240
QY	241	GGTGTGCTGCAAGCAAGTCTCAATCTCAATGCTGGAGAAATCCTCCAAATGTTGG	300
DB	241	GGTGTGCTGCAAGCAAGTCTCAATCTCAATGCTGGAGAAATCCTCCAAATGTTGG	300
QY	301	GAAGATGTTTTCGCTTTTGGCAGAAATCTGTTATGATACAAATCTTCGTTGCTGG	360
DB	301	GAAGATGTTTTCGCTTTTGGCAGAAATCTGTTATGATACAAATCTTCGTTGCTGG	360
QY	361	CTCTAATGTCAGAGATTCTACAGAACCTTGATGCTCTGCAGACCACTTGTCTACCAT	420
DB	361	CTCTAATGTCAGAGATTCTACAGAACCTTGATGCTCTGCAGACCACTTGTCTACCAT	420
QY	421	CTACCCAGGAATCGCTGCACCTTCTTTAGTGCATGATGATGAGAAAGGCAAGGACT	480
DB	421	CTACCCAGGAATCGCTGCACCTTCTTTAGTGCATGATGATGAGAAAGGCAAGGACT	480
QY	481	CATTGTGCTACTACTACAGAGAGAGGACTTCCAGGATATGTCATTTGGAATCATCAA	540
DB	481	CATTGTGCTACTACTACAGAGAGAGGACTTCCAGGATATGTCATTTGGAATCATCAA	540
QY	541	AACAGTGGCACAACAAATCCATGGCACTGAAATAGACATGAGGTTATTTCAGCAAGAA	600
DB	541	AACAGTGGCACAACAAATCCATGGCACTGAAATAGACATGAGGTTATTTCAGCAAGAA	600
QY	601	TGAAGAATGTGATCATCTCAATTTTAAATGAAGAAAGAGTCAAAAGAGAGGATT	660
DB	601	TGAAGAATGTGATCATCTCAATTTTAAATGAAGAAAGAGTCAAAAGAGAGGATT	660
QY	661	TTATGAAGATCTTGACAGATTGAGAAATGGTACCCAGGAATCACGCATCAGGCCATA	720
DB	661	TTATGAAGATCTTGACAGATTGAGAAATGGTACCCAGGAATCACGCATCAGGCCATA	720
QY	721	TACATTTCTCAAGCTTTTCTTTTCATATATATTTGACCGGACCTAGTGTCTACTCA	780
DB	721	TACATTTCTCAAGCTTTTCTTTTCATATATATTTGACCGGACCTAGTGTCTACTCA	780

DB	721	TACATTTCTCAAGCTTTTCTTTTCATATATATTTGACCGGACCTAGTGTCTACTCA	780
QY	781	GTGTGGCAATGCTATATACAGAGTTCTCCCCAGCTCCAGCTGGGAATTTGAGCCTTCT	840
DB	781	GTGTGGCAATGCTATATACAGAGTTCTCCCCAGCTCCAGCTGGGAATTTGAGCCTTCT	840
QY	841	GTCTGTCTTCTCGCTGGTTCGCTCATATATTGATATTAGTTTCCATGGGATCCTTTCTCA	900
DB	841	GTCTGTCTTCTCGCTGGTTCGCTCATATATTGATATTAGTTTCCATGGGATCCTTTCTCA	900
QY	901	CATCAATACTGTTTTTGTATTGAGAAGCAAGAGATTGTTGGATGTGGAGAAATTAGA	960
DB	901	CATCAATACTGTTTTTGTATTGAGAAGCAAGAGATTGTTGGATGTGGAGAAATTAGA	960
QY	961	ATGTGAGGATGAACCTGACTGGGACTGAGATCAGCTGTCTACCTCTCAAGGGTCAATGAT	1020
DB	961	ATGTGAGGATGAACCTGACTGGGACTGAGATCAGCTGTCTACCTCTCAAGGGTCAATGAT	1020
QY	1021	CTACTTACCTGAAGCAGATAGCATATCTTTTCTATTTTCCCAAGTGTCTATGAACCTGA	1080
DB	1021	CTACTTACCTGAAGCAGATAGCATATCTTTTCTATTTTCCCAAGTGTCTATGAACCTGA	1080
QY	1081	CGATTTGACAAGGAGGGCTGTATCTAAGTACATCCCTCTGCATGATGCCACGGCCA	1140
DB	1081	CGATTTGACAAGGAGGGCTGTATCTAAGTACATCCCTCTGCATGATGCCACGGCCA	1140
QY	1141	TCCTGTTCTTTTGGGAGAACAAATTTAGAGAGGAATACAAACTCACCAAGAACTGGAAT	1200
DB	1141	TCCTGTTCTTTTGGGAGAACAAATTTAGAGAGGAATACAAACTCACCAAGAACTGGAAT	1200
QY	1201	CCTCACTGACAGGCTACAGCTCACGTTAAGAGCCCTGGAAGATGAAAAGAAAGACAGA	1260
DB	1201	CCTCACTGACAGGCTACAGCTCACGTTAAGAGCCCTGGAAGATGAAAAGAAAGACAGA	1260
QY	1261	CACATTGCTGTAATCTGCTCTTCTCCGCTCTGTGCCAATGAGCTGCGCACAAAGGCTCC	1320
DB	1261	CACATTGCTGTAATCTGCTCTTCTCCGCTCTGTGCCAATGAGCTGCGCACAAAGGCTCC	1320
QY	1321	AGTGCTGCGCAAAAGATATGACAAATGTGACCACTCTCTTTAGTGGCATTGTGGCTTCAA	1380
DB	1321	AGTGCTGCGCAAAAGATATGACAAATGTGACCACTCTCTTTAGTGGCATTGTGGCTTCAA	1380
QY	1381	TGCTTTCTGTAGCAAGATGCAATCTGGAGAAAGAGCCATGGAAGTGTCAACTCTCTCAA	1440
DB	1381	TGCTTTCTGTAGCAAGATGCAATCTGGAGAAAGAGCCATGGAAGTGTCAACTCTCTCAA	1440
QY	1441	CGACCTCTACACAGATTGACACACTGACTGATTTCCCGGAAAACCCCAATTTGTTTATAA	1500
DB	1441	CGACCTCTACACAGATTGACACACTGACTGATTTCCCGGAAAACCCCAATTTGTTTATAA	1500
QY	1501	GGTGAGAGCTGTTGGTGACAAAGTATATGACAGTGTGTTTACCAGAGCCATGCAATCA	1560
DB	1501	GGTGAGAGCTGTTGGTGACAAAGTATATGACAGTGTGTTTACCAGAGCCATGCAATCA	1560
QY	1561	CCATGCAAGATCCATCTGCCACCTGGCCTTGGACATGATGGAATTTGTGGCCAGGTTCA	1620
DB	1561	CCATGCAAGATCCATCTGCCACCTGGCCTTGGACATGATGGAATTTGTGGCCAGGTTCA	1620
QY	1621	AGTAGATGTTGAATCTGTTTACATAACATAGGATACACACTGAGAGGTTAGTTACAGG	1680
DB	1621	AGTAGATGTTGAATCTGTTTACATAACATAGGATACACACTGAGAGGTTAGTTACAGG	1680
QY	1681	TGTCATAGGACAGCGGATGCTCTGATCTGTTTGGGAATCTGTCAACTCACAAG	1740
DB	1681	TGTCATAGGACAGCGGATGCTCTGATCTGTTTGGGAATCTGTCAACTCACAAG	1740
QY	1741	CCGAACAGAAACCCAGAGAGAAAGGAAAAATAAATGTGTCTGAATATACATACAGATG	1800
DB	1741	CCGAACAGAAACCCAGAGAGAAAGGAAAAATAAATGTGTCTGAATATACATACAGATG	1800
QY	1801	TCCTTATGCTCCAGAAAAATTCAGATCCCAATTTCACTTTGGAGCACAGAGGCCAGTGTC	1860
DB	1801	TCCTTATGCTCCAGAAAAATTCAGATCCCAATTTCACTTTGGAGCACAGAGGCCAGTGTC	1860

QY 730 CAAAGCTTTTCCCTTTCATATATATTTGACCGGGACCTAGTGGTCACTCAGTGTGGCAA 789
Db 2423 CAAAGCTTTTCCCTTTCATATATATTTGACCGGGACCTAGTGGTCACTCAGTGTGGCAA 2364
QY 790 TGCATATACAGAGTTCTCCCGCAGCTCCAGCCTGGGAATTCAGCCTTCTGCTCTT 849
Db 2363 TGCATATACAGAGTTCTCCCGCAGCTCCAGCCTGGGAATTCAGCCTTCTGCTCTT 2304
QY 850 CTCGCTGGTTCGCTCCTATATATAGTTTCCATGGGATCCTTCTCACATCAATAC 909
Db 2303 CTCGCTGGTTCGCTCCTATATATAGTTTCCATGGGATCCTTCTCACATCAATAC 2244
QY 910 TGTGTTTGTATGAGAAGCAAGAGGATGTTGGATGTTGGAGAAATTAAGATGTGAGGA 969
Db 2243 TGTGTTTGTATGAGAAGCAAGAGGATGTTGGATGTTGGAGAAATTAAGATGTGAGGA 2184
QY 970 TGAACCTGACTGGGACTGAGATCAGCTGCTTAC-GTCTCAAGGTCGAATGATCTACTTAC 1028
Db 2183 TGAACCTGACTGGGACTGAGATCAGCTGCTTACAGTCTCAAGGTCGAATGATCTACTTAC 2124
QY 1029 GTGAACGAGATAGCATCTTTTCTATGTTTCAACAGTGTCTGATGAACCTGGAGGATTGGA 1088
Db 2123 GTGAACGAGATAGCATCTTTTCTATGTTTCAACAGTGTCTGATGAACCTGGAGGATTGGA 2064
QY 1089 CAAGGAGGGGTGTATCTAAGTGACATCCCTCTGCATGATGCCAGCGCGATCTTTGTTTC 1148
Db 2063 CAAGGAGGGGTGTATCTAAGTGACATCCCTCTGCATGATGCCAGCGCGATCTTTGTTTC 2004
QY 1149 TTTTGGGAGACAAATTTAGAGGAAATACAACTCACCAAGACTGGAATCCCTCACCTG 1208
Db 2003 TTTTGGGAGACAAATTTAGAGGAAATACAACTCACCAAGACTGGAATCCCTCACCTG 1944
QY 1209 ACAGGCTACAGCTCAGCTTAAGAGCCCTGGAAGATGAAAGAAAGACAGACACATTCG 1268
Db 1943 ACAGGCTACAGCTCAGCTTAAGAGCCCTGGAAGATGAAAGAAAGACAGACATTCG 1884
QY 1269 TGTATTCTGCTTCTCCGCTGTTGGCCAAATGAGCTGGGACAAAGCGTCCAGTGCCTG 1328
Db 1883 TGTATTCTGCTTCTCCGCTGTTGGCCAAATGAGCTGGGACAAAGCGTCCAGTGCCTG 1824
QY 1329 CCAAAAGATATGACAATGTGACCATCTCTTTAGTGGCAATGTGGGCTTCAATGCTTCT 1388
Db 1823 CCAAAAGATATGACAATGTGACCATCTCTTTAGTGGCAATGTGGGCTTCAATGCTTCT 1764
QY 1389 GTAGCAAGCATGATCTGGAGAGGAGCCATGAAGATCGTCAACCTCTCAACGACCTCT 1448
Db 1763 GTAGCAAGCATGATCTGGAGAGGAGCCATGAAGATCGTCAACCTCTCAACGACCTCT 1704
QY 1449 ACACCAAGATTGACACACTGACTGATCCCGGAAAACCCATTTGTTATNAGGTGGAGA 1508
Db 1703 ACACCAAGATTGACACACTGACTGATCCCGGAAAACCCATTTGTTATNAGGTGGAGA 1644
QY 1509 CTGTTGGTGACAGTATATGACAGTGAAGTGGTTTACGAGAGCATGATTCACCATGAC 1568
Db 1643 CTGTTGGTGACAGTATATGACAGTGAAGTGGTTTACGAGAGCATGATTCACCATGAC 1584
QY 1569 GATCCATCTGCCACCTGGCCTTGGACATGATGGAATTCCTGGCCAGGTTCAAGTAGATG 1628
Db 1583 GATCCATCTGCCACCTGGCCTTGGACATGATGGAATTCCTGGCCAGGTTCAAGTAGATG 1524
QY 1629 GTCAATCTGTCAGATAACAATAGGATACACACTGGAGAGTAGTTACAGTGTGCATAG 1688
Db 1523 GTCAATCTGTCAGATAACAATAGGATACACACTGGAGAGTAGTTACAGTGTGCATAG 1464
QY 1689 GACAGGGATGCTCGACTACTGCTTTTGGGAATCTGTCAACCTCAACAAGCCGAACAG 1748
Db 1463 GACAGGGATGCTCGACTACTGCTTTTGGGAATCTGTCAACCTCAACAAGCCGAACAG 1404
QY 1749 AAACCAAGAGAGAAAGGAAATTAATGTCTGTGAATATACATACAGATGCTTATGT 1808
Db 1403 AAACCAAGAGAGAAAGGAAATTAATGTCTGTGAATATACATACAGATGCTTATGT 1344
QY 1809 CTCGAGAAATTCAGATCCACAATTCCTCACTTGGAGCAGAGAGCCGAGTGTCCATGAAG 1868

Db 1343 CTCGAGAAATTCAGATCCACAATTCCTCACTTGGAGCAGAGGCCAGTGTCCATGAAG 1284
QY 1869 GCARAAAGAACCAATGCAAGTTTGGTTTCTATCCAGAAAAAATACAGAGACAGAGAA 1928
Db 1283 GCARAAAGAACCAATGCAAGTTTGGTTTCTATCCAGAAAAAATACAGAGACAGAGAA 1224
QY 1929 CAAAGCAGGATGATGACTGAATCTTGGATTATGGGCTGAAGAGAGTACAGACTAGGTTTC 1988
Db 1223 CAAAGCAGGATGATGACTGAATCTTGGATTATGGGCTGAAGAGAGTACAGACTAGGTTTC 1164
QY 1989 CAGTTTCTCCTAACACGTCGCAAGCCAGGAGCAGTTCTTCCCTATGATACAGATTTT 2048
Db 1163 CAGTTTCTCCTAACACGTCGCAAGCCAGGAGCAGTTCTTCCCTATGATACAGATTTT 1104
QY 2049 CTTTGTCTCTTGTCCATACCCCAAGACTTCTCTAGATATATCTCTCACTATCCGTTA 2108
Db 1103 CTTTGTCTCTTGTCCATACCCCAAGACTTCTCTAGATATATCTCTCACTATCCGTTA 1044
QY 2109 TTTCAACCTTAGCTCTGCTTCTTATTAATCTTGTAGGCTTTAGTATATATCTAAAGTTGG 2168
Db 1043 TTTCAACCTTAGCTCTGCTTCTTATTAATCTTGTAGGCTTTAGTATATATCTAAAGTTGG 984
QY 2169 CTTTGTATGATGATGATGAGCTTCATGCTTCTTAAATCTACTACAAGCATTAACCTAA 2228
Db 983 CTTTGTATGATGATGATGAGCTTCATGCTTCTTAAATCTACTACAAGCATTAACCTAA 924
QY 2229 CATGCTGATCTGCAAGTAGTAGGACCCCAATAAATATTTGTTGAATTTAGTTAAATGAA 2288
Db 923 CATGCTGATCTGCAAGTAGTAGGACCCCAATAAATATTTGTTGAATTTAGTTAAATGAA 864
QY 2289 CTGAACAGTGTGTCATGCTATATATATATCATGTTTACCAATCTGTTAGTGTTC 2348
Db 863 CTGAACAGTGTGTCATGCTATATATATATCATGTTTACCAATCTGTTAGTGTTC 804
QY 2349 CACATATATGATATGATATATTTTAAGACTATAATGATTAAGTTTATATCATGTTGG 2408
Db 803 CACATATATGATATGATATATTTTAAGACTATAATGATTAAGTTTATATCATGTTGG 744
QY 2409 TGTATATCATTTAGAAATCATTTCTTAAAGGAGT 2443
Db 743 TGTATATCATTTAGAAATCATTTCTTAAAGGAGT 709

RESULT 3
AAH33895
ID AAH33895 standard; cDNA; 1302 BP.
XX
AC AAH33895;
XX
DT 03-SEP-2001 (first entry)
XX Human colon cancer antigen encoding cDNA SEQ ID NO:951.
DE Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 4; ss.
KW
XX Homo sapiens.
XX
XX WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
DR

QY 250 TGCAGCAAGTCCCTCAATCTCAATGCTGGAGAAATCCCTCCAAATGTTGGGAAGATGT 309
 |||||
 DB 264 TGCAGCAAGTCCCTCAATCTCAATGCTGGAGAAATCCCTCCAAATGTTGGGAAGATGT 323
 |||||
 QY 310 TTTCGTCCTTTGCCAAGATCTGGTATATACATCAATCTTGGCTGCTGGCTCTAATGT 369
 |||||
 DB 324 TTTCGTCCTTTGCCAAGATCTGGTATATACATCAATCTTGGCTGCTGGCTCTAATGT 383
 |||||
 QY 370 CAGAGAAATCTACAGAACTTGATGCTCTGCAGCAGCACCTTGCTTACCATCTACCCAGG 429
 |||||
 DB 384 CAGAGAAATCTACAGAACTTGATGCTCTGCAGCAGCACCTTGCTTACCATCTACCCAGG 443
 |||||
 QY 430 AATGCGTGCACTTCC 445
 |||||
 DB 444 AATGCGTGCACTTCC 459

RESULT 5

ABL02665
 ID ABL02665 standard; cDNA; 2850 BP.

XX
 AC ABL02665;
 XX

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 2477.

DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB58562.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Claim 1; SEQ ID NO 2477; 2lpp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2850 BP; 734 A; 756 C; 788 G; 572 T; 0 other;

Query Match

Best Local Similarity 15.2%; Score 371.4; DB 23; Length 2850;

Matches 694; Conservative 58.1%; Pred. No. 4.4e-86;

Mismatches 0; Mismatches 471; Indels 30; Gaps 1;

QY 705 CACGCAATCAGCCCATATACATCTCTGCAAAAGCTTTTCTTTTCATATAATATTGACCGGG 764
 |||||
 DB 1276 CACTCATATCGCCGCGCACTTTCTGCAAGGTGTCCCTTCCACCTGATGTTTCGATAGGC 1335
 |||||
 QY 765 ACCTAGTGTCTCAGTGTGGCAATGCTATATACAGAGTTCTCCCGCAGCTCCACGCTG 824
 |||||
 DB 1336 AATGAAATCTGTTAGCGCCGCAAGGCTGTCTCGAGTAATTTCCAGAGTTGCCGAGG 1395
 |||||
 QY 825 GGAATTTGACGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884
 |||||
 DB 1396 AGACTGTTCCCTAATAGAGTGGTGGAGCCATCCGCGCGACCTGCGAGCTCAACTTCG 1455
 |||||
 QY 885 ATGGGATCCTTTCTCATCAATACATCTGTTTGTATTGAGAGCAAGAGGATTTGTTGG 944
 |||||
 DB 1456 AGAATCATCTGTCCTCATATAACACCATCTATGCTCTGCAGACACCTCAGGGGGCCATGA 1515
 |||||
 QY 945 ATGTGGAGAAATAGNATGTGAGGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1004
 |||||
 DB 1516 GC-----AGCGCGCAGCAGCAGCGCTTTTCTGAGAC 1545
 |||||
 QY 1005 TCAAGGGTCAAAATGATCTACTTTACCTGAAGCAGATAGCATACTTTTCTATGTTACCAA 1064
 |||||
 DB 1546 TGAAGGGCAATGATGATACATTCGGAAACGGATCGGATCTGTTCAATGCTATCCCA 1605
 |||||
 QY 1065 GTGTCATGAACCTGGACGATTTGACAAGGAGAGGCTGTATCTAAGTGAATCCCTCTGC 1124
 |||||
 DB 1606 GTGTCATGAATCTGGATGATCTGACCAAGAAGGACTTACATCTCGGATGTGCCCTGC 1665
 |||||
 QY 1125 ATGATGCCACGCGCATCTGTTCTTTTGGGAGACAATTTAGAGGAGATACAACTCA 1184
 |||||
 DB 1666 ACAGCGTCCAGGGATTTGGTCTCTCCGAGAAAGTTCGAGGGCGGAGTACAAGCTCA 1725
 |||||
 QY 1185 CCAAGAACTGGAAATCTCACTGACGAGCTACAGCTACGTTAAGAGCCCTGGAAAGATG 1244
 |||||
 DB 1726 CCAAGAACTGGAAATGCTCACCAGCAAGCTCAGCAGACATTCGCGATCTGGAGAGC 1785
 |||||
 QY 1245 AAAAGAAAAGACAGACATTTGCTGTATCTGCTCTCTCCGCTGTTGGCAATGAGC 1304
 |||||
 DB 1786 AGAAGCAGAAGACCGACAGGCTGCTCTATTCGGTGTGCGCAAGTCCGTGGCAATGAGT 1845
 |||||
 QY 1305 TCGGCAACAAGCGTCCAGTGCCTGCCAAAAGATATGACAATGTGACCATCCCTTTAGTG 1364
 |||||
 DB 1846 TACGACATCAGCGTCTCTGCGCCCAACGCTACGACTCCGCTGAGCTTGATGTTCTCCG 1905
 |||||
 QY 1365 GCATTGTGGGCTTCAATGCTTCTGTAGCAAGCATGCTCTGGAGAGGAGGACCATGAAGA 1424
 |||||
 DB 1906 GCATCGTGGCTTTGGCCAAATACTGTGCGGCCCAACACAGATCCCGAGCGGCGCATGAAA 1965
 |||||
 QY 1425 TCGTCAACCTCTCAACGACCTCTACACCAGATTTGACACACTGACTGATTCGCGGAAA 1484
 |||||
 DB 1966 TCGTGAATAATGCTGAATGAATCTACACGCTCTCTGATGCTCTGACCGACTCCAAGCGAA 2025
 |||||
 QY 1485 ACCCATTTGTTTATAGGTGGAGACTGTTGGTGACAAGTATATGACAGTGAAGTGTTCAC 1544
 |||||
 DB 2026 ACCTGAACGTGTACAGGTGTTGAACAGTTGGCGATAAGTACATGGCCGTTTCGGGACTGC 2085
 |||||
 QY 1545 CAGAGCCATGCAATTCACCATGCAACGATCCATCTGCCACCTGGCCTTGGACATGAGGAAA 1604
 |||||
 DB 2086 CCGATCACTCGGAGGATCATGCCAAGTGCATGCCAGAGTGGCCCTCGATATGATGACGA 2145
 |||||
 QY 1605 TTGCTGGCCAGGTTCAAGTAGATGGTGAATCTGTTAGATAACAATAGGAGTACACACTG 1664
 |||||
 DB 2146 TGGCCCAAGAACCTCAAAATGGGATCCAAATCCAGTGCAAATCACCATTAGGACATTCGG 2205
 |||||
 QY 1665 GAGAGGTAGTTACAGGTGTCATAGACAGCGGATGCTCGATGCTCTCTTTTGGGAATA 1724
 |||||
 DB 2206 GCGAAGTGTGACTGGAGTGTGCGCAACCGAGTGCACGCTACTGCTGTGTCGGAATA 2265
 |||||
 QY 1725 CTGTCAACCTCAAGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1784
 |||||
 DB 2266 CTGTCAATCTACAGCGCGGAGACCACTGGGCTTCCGGGGCGGCACTCAAGCTCAGCG 2325
 |||||
 QY 1785 AATATACATACAGATGCTTTATGCTCTCCAGAAAATTCAGATCCACAATTCACATCCAGG 1844
 |||||

Db 2326 AAGAACCTATCGGCTCTCTGTATGGCGATCAACAGGATGATTCGTTCCACTTGGAGT 2385
QY 1845 ACAGAGGCCAGTCTCCATGAAGGCAAAAGAACCAATGCAAGTTTGGTTCT 1899
Db 2386 ACCGGGACCCCTTATCATGAAGGGCAACACCGCCGACGCTGTTGGTTCT 2440

RESULT 6
ID AAS94827 standard; DNA; 2954 BP.
XX AAS94827;
AC
XX
DT 14-FEB-2002 (first entry)
XX
DE Human DNA sequence #82 expressed during foam cell differentiation.
XX
KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200177389-A2.
XX
PD 18-OCT-2001.
XX
PF 04-APR-2001; 2001WO-US111128.
XX
PR 05-APR-2000; 2000US-195106P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI Tai J;
PI WPI; 2002-010925/01.
XX
PT Composition useful for diagnosis of conditions, disorders or diseases
PT associated with atherosclerosis, comprises several polynucleotides that
PT are differentially expressed in foam cell development -
XX
PS Claim 1; Page 134-135; 315pp; English.
XX
CC The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used
CC as PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation.
XX
SQ Sequence 2954 BP; 735 A; 779 C; 760 G; 680 T; 0 other;

Query Match 8.2%; Score 200.2; DB 24; Length 2954;
Best Local Similarity 55.8%; Pred. No. 1.7e-41;
Matches 457; Conservative 0; Mismatches 338; Indels 24; Gaps 3;

QY 1005 TCAAGGTCGAATGATCTACTACCTGAAGCAGATAGCATCTTTTCTATGTTCCACAA 1064
Db 1607 TCAAGGCAAAATGATCCATGTTCCAGATCAATTCATTTATTTTGGGCTCTCCAT 1666
QY 1065 GTGTCATGAACCTGGAGCTTTGACAAAGGAGGCGGTGTATCTAAGTCACATCCCTCTGC 1124
Db 1667 GTGTGGCAAGATTGGATGAACATCATGGCGGAGGGCTACATCTCTCAGACATCCCTATCC 1726
QY 1125 ATGATGCCACGGCGGATCTTGTCTTTTGGGAGAACAAATTAGAGAGGAATACAACTCA 1184

Db 1727 ATGATGCCACCGAGATGTCATTTTGGTTGGTGAGCAGCAAGGCCCAAGATGGGTTGA 1786
QY 1185 CCACAGAAGCTGGAAATCCCTCACTGACAGGCTACAGCTACAGTTAAGAGCCCTGGAGATG 1244
Db 1787 AGAAAAGGATGATAAATTAAGGCAACTTTAGAAAGAACTACACAGGCCCTGGAAAG 1846
QY 1245 AAAAGAAAAGACAGACACATTTCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1304
Db 1847 AGAAAAGGAGACAGTGGATCTTCTATATTTCTATTTTCTCTCTCTCTCTCTCTCTCT 1906
QY 1305 TCGGGCACAAGCGTCCAGTCCCTGCGCCAAAGATATGACAAATGAGACCATCTCTTTAGTG 1364
Db 1907 TATGGCAAGGGCAGCAAGTACAGGCCAGAAAGTTTGTATGATGTGCACCTGCTCTTTTCAG 1966
QY 1365 GCATTTGTGGCTTCAATGCTTTCTGTAGCAAGCATCATCTGAGGAAGAGGCCATCAAGA 1424
Db 1967 ACATTTGTGGCTTCAAGCCATATGTCGCCAGTGTAC-----TCCATCGAAG 2014
QY 1425 TCGTCAACCTCTCAACAGCTCTACACAGATTTTGACACACTGATGATTCGCCGAAA 1484
Db 2015 TAATCAGCATGCTGAATGAATGTACACCAAGATTTGACACCACTGTTGGATTTTGGAA-- 2072
QY 1485 ACCATTTGTTTATAGGTGGAGACTGTTGGTGACAAGTATATGACAGTGTGTTTAC 1544
Db 2073 -----TATTTATAGGTGGAACAATAGTGTCTACTGTGTTCAGCAGGGCTCC 2125
QY 1545 CAGAGCATGCTTACCATGACAGTCCATCTGCCACCTGGCTTGGACATGATGAAA 1604
Db 2126 ACAGAAAAGCTCTGCCATGTAAACCCATGCTGTGATGGCTTGAAGATGATGAAAC 2185
QY 1605 TTGCTGGCAGGTTCAAGTA---GATGGTGAATCTCTTCTCAGATAACAAATAGGATACACA 1661
Db 2186 TTTTCAAGAGGTGCTGACACCTGATGGAAGACCGATTCAGATGAGGATGAGTAATCACT 2245
QY 1662 CTGGAGAGTGTACAGGTGTATAGGACAGCGGATGCTCTGATCTCTTTTGGGA 1721
Db 2246 CAGGCTCCCTGTGGCTGGATGTTGGGGTGGCAATGCGACCTTATTCCTTTTGGAA 2305
QY 1722 ATACTCTCACTCACAAGCCGAACCAAGAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1781
Db 2306 ATAAATGTCACACTGGCAAGCAATTCGAGTCGGGAAGTCACTCCCTCGGCGCATCAATGCA 2365
QY 1782 CTGAATATACATACAGATGCTTTATGCTCTCCAGAAAATT 1820
Db 2366 GCCCAACCACTTACCAATTTATAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2404

RESULT 7
ID AAZ51684 standard; cDNA; 2715 BP.
XX AAZ51684;
AC AAZ51684;
XX
DT 04-JUL-2000 (first entry)
XX
DE Human cyclic nucleotide-associated protein-3 (CNAP-3) cDNA.
XX
KW Cyclic nucleotide-associated protein-3; CNAP-3; human; cytosolic;
KW anti-arteriosclerotic; hepatotropic; anti-leukemic; anti-inflammatory;
KW immunomodulatory; anti-asthmatic; anti-anemic; anti-diabetic; diagnosis;
KW anti-sclerotic; dermatological; neuroprotective; anti-epileptic; cancer;
KW anti-Alzheimer's; anti-Parkinsonian; cerebroprotective; ophthalmological;
KW anti-fertility; anti-allergic; vasotropic; immunosuppressive;
KW hypotensive; gene therapy; prevention; treatment; arteriosclerosis;
KW cell proliferative disorder; autoimmune/inflammatory; diabetes mellitus;
KW neurological; vision; reproductive; smooth muscle; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 440..2512
FT /*tag= a

/product= "Human CNAP-3 protein"
 /note= "Shares 89% identity to human soluble guanylate
 cyclase large subunit"
 misc_binding 863..892
 /*tag= b
 /bound_moiety= "Primer or Probe"
 /note= "Useful for amplification or hybridisation
 techniques"
 WO200014248-A1.
 16-MAR-2000.
 03-SEP-1999; 99WO-US20287.
 04-SEP-1998; 98US-0148904.
 (INCY-) INCYTE PHARM INC.
 Hillman JL, Yue H, Guegler KJ, Corley NC, Patterson C, Tang YT;
 WPI; 2000-256994/22.
 P-PSDB; AAY70475.
 Isolated cyclic nucleotide associated proteins useful for preventing,
 diagnosing and treating cell proliferative, autoimmune/inflammatory,
 neurological, vision, reproductive and smooth muscle disorders -
 Example 3; Page 71-72; 78pp; English.
 The present sequence is the cDNA encoding human cyclic nucleotide
 associated protein-3 (CNAP-3), identified in incyte clone 159278,
 that is isolated from ADENIR01 cDNA library. It is expressed in
 nervous, reproductive, cardiovascular and developmental tissues.
 CNAP sequences may be used for prevention, treatment and diagnosis of
 diseases associated with altered CNAP expression such as, cell
 proliferative disorders (e.g. arteriosclerosis, cirrhosis, leukaemia,
 lymphoma and cancer of the breast, prostate, lung and brain), autoimmune/
 inflammatory disorders (e.g. asthma, anaemia, diabetes mellitus, multiple
 sclerosis and psoriasis), neurological disorders (e.g. epilepsy,
 Alzheimer's/Parkinson's disease and strokes), vision disorders (e.g.
 conjunctivitis, glaucoma, cataracts and retinitis pigmentosa),
 reproductive disorders (e.g. infertility, uterine fibroids, ectopic
 pregnancies and impotence) and smooth muscle disorders (e.g. angina,
 anaphylactic shock, Kearns-Sayre syndrome and hypertension). It can also
 be used for gene therapy.
 Sequence 2715 BP; 799 A; 604 C; 662 G; 650 T; 0 other;
 Query Match 7.98; Score 192.6; DB 21; Length 2715;
 Best Local Similarity 55.5%; Pred. No. 1.6e-39;
 Matches 447; Conservative 0; Mismatches 334; Indels 24; Gaps 3;
 QY 1004 CTCAGGGTCAAAATGATCTACTACCTGACGAGATAGCATCTTTTCTATGTTACCA 1063
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1535 CTCAGGCCCAATGATCTACATGTTGATCCAGTCGTAATCTGTTTGGGGTACCC 1594
 QY 1064 AGTGTATGAACCTGGACGATTTGACAGGAGAGGGCTGTATCTAAAGTGACATCCCTCTG 1123
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1595 TGTGTGGACAGATTAGAAGATTTTACAGGACGAGGGCTCTACCTCTCAGACATCCCAATT 1654
 QY 1124 CATGATCCACGGCGGCTCTCTTTTGGGAGAACAAATTTAGAGGAGTAACAATC 1183
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1655 CACAATGCTAGGGATGGTCTTAAATAGGGGAACAGCCCGAGCTCAAGATGGCGTG 1714
 QY 1184 ACCCAAGACTGGAATPCCCTCACTGACAGCTACAGCTTAAAGAGCCCTGGAGAT 1243
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1715 AAGAGAGGCTGGGAGCTGAGGCTACCTTTGAGCAAGCCCAAGCCCTGGAGGAG 1774
 QY 1244 GAAAGAAAAGACAGACATCTGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1303
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1775 GAGAGAAAAGACAGTAGACCTTCTGTGCTCCATATTTTCCCTGTGAGGTTGCTCAGCAG 1834

QY 1304 CTGGGCAAGCGTCCAGTCCCTGCGCAAGAGATATGACAAATGTGACCATCTCTTTAGT 1363
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1835 CTGTGGCAAGGCGAAGTTGTGCAAGCGCAAGAGTTCAATGATGTCACCATGCTCTCTCA 1894
 QY 1364 GGCATTGTGGCTCAATGCTTTCTGTAGCAAGCATGCTCTGGAGAGGAGCCATGAAG 1423
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1895 GACATGGTTGGGTTCATGCTCCATCTGCTCCAGTGGTC-----ACCGCTCAG 1942
 QY 1424 ATGCTCAACCTCCTCAACGACCTCTACACAGACATTTGACACACATGATGATTCGCCGAAA 1483
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1943 GTCATCACCATGCTCAATGCACTGTACACTCGCTTCGCCAGCAGTGTGGAGAGCTGA- 2001
 QY 1484 AACCATTTCTTTATAAGGTGGAGACTGTGGTGACAAGTATATGACAGTGAAGTGTGTTA 1543
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 2002 -----TGTCTACAGGTGGAGACCATTTGGGATGCTTATTGTGTAGCTGGGGATTA 2053
 QY 1544 CCAGAGCCATGCTACCATGCAAGCATCTGCCACCTGCGCTTGGACATGATGAA 1603
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 2054 CACAAGAGAGTGTACTCATGCTGTTTCAGATAGCGCTGATGCCCTGGAAGATGATGGAG 2113
 QY 1604 ATTGCTGGCCAGGTT---CAAGTAGATGTTGAATCTGTTTCAGATAACAATAGGATACAC 1660
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 2114 CTCTCTGATGAAGTTATGCTCTCCCATGGAGAACCTATCAAGATGCGAATGGACTGCAC 2173
 QY 1661 ACTGGAGAGTGTACAGTGTCTATGAGACAGGGATGCTCGTACTGCTCTTTTGGG 1720
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 2174 TCTGGATCAGTTTTTGTGCTGGCTGTTGAGTTAAATGCGCGTTACTGCTCTTTTGGGA 2233
 QY 1721 AATACTGTCAACCTCACAGCCGCAACAGAACACAGAGGAGGAGGAGGAGGAGGAGGAG 1780
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 2234 AACATGCTACTCTGGCTAACAAATTTGAGTCTCTGAGTGTACCACGAGGAGGAGGAGG 2293
 QY 1781 TCTGAATATATACAGATGCTCTTA 1805
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 2294 AGCCCAACAACCTTACAGATTACTCA 2318
 RESULT 8
 AAZ88938
 ID AAZ88938 standard; DNA: 3015 BP.
 AC AAZ88938;
 XX
 DT 26-MAY-2000 (first entry)
 XX
 DE Human soluble guanylylcyclase alphas DNA.
 XX
 KW Human; guanylylcyclase alphas; hscGcalpha1; hscGbeta1; soluble;
 KW guanylylcyclase beta1; antiarteriosclerotic; vasotropic; hypotensive;
 KW gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;
 -KW peripheral arterial occlusive disease; arterial hypertension; ds.
 XX
 OS Homo sapiens.
 XX
 PN DE19837015-A1.
 XX
 PD 24-FEB-2000.
 XX
 PF 14-AUG-1998; 98DE-1037015.
 XX
 PR 14-AUG-1998; 98DE-1037015.
 XX
 PA (VASO-) VASOPHARM BIOTECH GMBH & CO KG.
 XX
 PI Schmidt H, Zabel U, Poller W;
 DR WPI; 2000-184044/17.
 DR P-PSDB; AAY51607.
 XX
 PT New human soluble guanylate cyclase alphas/beta1 and the nucleic acid
 PT encoding the subunits, useful for producing diagnostic antibodies, and
 PT for somatic gene therapy of arteriosclerosis -
 XX

XX	16-SEP-2002 (first entry)	
XX	Human prostate expression marker cDNA 27559.	
XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;	
XX	pharmacogenomic marker; gene; ss.	
XX	Homo sapiens.	
XX	WO200160860-A2.	
XX	23-AUG-2001.	
XX	20-FEB-2001; 2001WO-US05171.	
XX	17-FEB-2000; 2000US-183319P.	
XX	16-MAR-2000; 2000US-189862P.	
XX	23-MAY-2000; 2000US-207454P.	
XX	09-JUN-2000; 2000US-211314P.	
XX	18-JUL-2000; 2000US-219007P.	
XX	13-DEC-2000; 2000US-255281P.	
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
XX	Schlegel R, Endege WO, Monahan JE;	
XX	WPI; 2001-662795/76.	
XX	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -	
XX	Claim 1; Page 5633-5634; 11750pp; English.	
XX	The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:	
XX	(a) assessing whether a patient is afflicted with prostate cancer;	
XX	(b) monitoring the progression of prostate cancer in a patient;	
XX	(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;	
XX	(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;	
XX	(e) selecting a composition for inhibiting prostate cancer in a patient;	
XX	(f) assessing the prostate cell carcinogenic potential of a compound;	
XX	(g) determining whether prostate cancer has metastasized in a patient;	
XX	(h) assessing the aggressiveness or indolence of prostate cancer in a patient;	
XX	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.	
XX	Sequence 4742 BP; 1450 A; 933 C; 998 G; 1355 T; 6 other;	
XX	Query Match 7.88; Score 191; DB 23; Length 4742;	
XX	Best Local Similarity 55.4%; Pred. No. 5.5e-39;	
XX	Matches 446; Conservative 0; Mismatches 335; Indels 24; Gaps 3;	
QY	1004 CTCAGGGTCAAAATGATCTACTTTACTCTGAAGCAGATAGCATCTTTTCTCTTCACCA 1063	
Db	1618 CTCAGGCCAAATGATCTACATGTTGAATCAGTGCATCTGTTTTGGGGTGCACC 1677	
QY	1064 AGTGTGATCAACCTGGACGATTTGACAAGGAGAGGGCTGTATCTAAGTCACATCCCTCTG 1123	
Db	1678 TGTGTGCACAGATTAGAAGATTTTACAGGACGAGGCTCTACCTCTCAGACATCCCAATT 1737	
QY	1124 CATGATGCCACGCCGATCTGTTCTTTTGGGAGAACAAATTAGAGAGGAATACAACCTC 1183	
Db	1738 CACAATGCTACTGAGGATGTGGTCTTAATAGGGGAACAGCCCGAGCTCAAGATGGCCTG 1797	
QY	1184 ACCCAAGACTGGAATCCCTCACTGACAGGCTACGCTACGTTAAGAGCCCTGGAGAT 1243	
Db	1798 AAGAGAGGCTGGGAGGCTGAAGGCTTACCCCTTGAGCAAGCCCAAGCCCTGGAGGAG 1857	

(EOSB-) EOS BIOTECHNOLOGY INC.

PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX WPI; 2002-471335/50.

DR P-PSDB; ABG61878.

XX Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes
PT are expressed in a prostate tissue -

PS Claim 22; Page 362-363; 436pp; English.

XX The present invention relates to methods of detecting a prostate
CC cancer-associated transcript in a cell from a patient. The method
CC comprises contacting a biological sample from the patient with
CC prostate cancer-associated polynucleotides (designated PC genes) that
CC selectively hybridize to a sequence that is at least 80% identical
CC to them. The prostate cancer-associated polynucleotide sequences
CC are differentially expressed in prostate tumour tissue or in
CC prostate cancer and are derived from the tissues of various
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC The methods of the invention are useful for diagnosing and treating
CC prostate cancer in mammals. The prostate cancer-associated genes are
CC useful for diagnosing or treating prostate cancer, as well as for
CC identifying modulators of prostate cancer or agents that inhibit
CC prostate cancer. The nucleic acid sequences are particularly useful
CC in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences.

SQ Sequence 3004 BP; 874 A; 679 C; 710 G; 741 T; 0 other;

Query Match 7.6%; Score 186.2; DB 24; Length 3004;
Best Local Similarity 55.0%; Pred. No. 7.8e-38;
Matches 443; Conservative 0; Mismatches 338; Indels 24; Gaps 3;

QY 1004 CTCAGGTCGAATGATCTACTTACCTGAAGCAGATACATACATCTTTTCTATGTTTCAACA 1063
DB 1609 CTCAGGTCGAATGATCTACTTACCTGAAGCAGATACATACATCTTTTCTATGTTTCAACA 1668
QY 1064 AGTGTCATGAACCTGGAGGATTTGACAGAGAGGGCTGTATCAAGTGACATCCCTCTG 1123
DB 1669 TGTGTGGACAGATTAGAAGATTTTACAGAGAGGGCTGTATCAAGTGACATCCCTCTG 1728
QY 1124 CATGATGCACGCGCGATCTTTCTTTTGGGAGAACAAATTTAGAGAGGAATACAAACTC 1183
DB 1729 CACAATGCACCTGAGGATGTGTCTTAATAGGGGAACAAGCCGAGCTCAAGATGGCCTG 1788
QY 1184 ACCCAGAACTGGAATCTCTACTGACAGGCTACAGCTCAGTTAAGAGCCCTGGAGAT 1243
DB 1789 AAGAAGAGGCTGGGGAAGCTGAAGGCTACCCCTTGAGCAAGCCCAAGCCCTGGAGAG 1848
QY 1244 GAAAGAAAGAGACAGACACATTTGCTGTATCTGCTTCTCTGCTGCTGCTGCTGCTGCTG 1303
DB 1849 GAGAGAAAGAGACAGTACAGCTTCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1908
QY 1304 CTGGGCAACAGCGCTCCAGTCCCTGCCAAAGATATGACATGTACCATCTCTTTTCTTCTT 1363
DB 1909 CTGTGGCAAGGCAAGTGTGCAAGCAAGAGTTCAGTAATGTACCATCTCTTCTTCTTCT 1968
QY 1364 GGCATGTGGGCTCAATGCTTCTGTAGCAAGCATGCTGGAAGAGGAGCCATGAAG 1423
DB 1969 GACATCGTGGGCTTCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2016
QY 1424 ATCGTCAACCTCCCTCAAGCCTCTACACAGATTTGACACATGCTGATTCCTCGGAA 1483
DB 2017 GTCAATCACCAGTCAATGCACTGTACACTGCTCTGACAGAGTGTGGAGAGCTGGA- 2075
QY 1484 AACCCATTTTGAAGGTTGGAGACTGTTGGTGCACAGATATATGACATGAGTGGTGGTTTA 1543
DB 2076 -----TGCTACAGGTGGAGACCATTCGGATGCCCTATTTGTTGGCTTGGGGGATTA 2127

QY 1544 CCAGAGCCATGCATTCACCATGACGATCCATCTGCGCACCTGGCCTTGGACATGATGAA 1603
DB 2128 CACAAAGAGAGTGTATCTACTGCTCTTCCAGATAGCGCTGATGCGCTCGAAGATGATGAG 2187
QY 1604 ATTGCTGCCAGGTTCAAGTA--GATGGTGAATCTGTTCCAGATAACAATAGGGATACAC 1660
DB 2188 CTCCTGTATGAGTATGCTCTCCCATGGAGAACCTATCAAGATCGGAATGGAGCTGCAC 2247
QY 1661 ACTGGAGAGGTAGTTACAGGTGTCTATAGGACAGCGGATGCCCTCGATACTGTCTTTTGGG 1720
DB 2248 TCTGATCAGTTTTTGTGCTGGCTGCTGTTGGAATTAATGCGCGTACTGCTCTTTTGGG 2307
QY 1721 AATCTGTCAACCTCACAAGCCGGAACACACAAACACAGAGAAAGGGAATAAATGTG 1780
DB 2308 AACATGTCTACTCTGGCTTAACAAATTTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2367
QY 1781 TCTGAATATACATACAGATGCTTTA 1805
DB 2368 AGCCCAACAACTTACAGATTACTCA 2392

RESULT 14

ABK92259

ID ABK92259 standard; DNA; 3004 BP.

AC ABK92259;

DT 15-AUG-2002 (first entry)

DE Prostate cancer-associated DNA sequence #145.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;

KW gene therapy; gene; ds.

OS Mammalia.

XX WO200230268-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US32045.

XX 13-OCT-2000; 2000US-0687576.

XX 08-DEC-2000; 2000US-0733288.

XX 24-JAN-2001; 2000US-0733742.

XX 16-MAR-2001; 2001US-263957P.

XX 16-MAR-2001; 2001US-276791P.

XX 06-APR-2001; 2001US-281922P.

XX 30-APR-2001; 2001US-286214P.

XX 04-MAY-2001; 2001US-288589P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX WPI; 2002-471335/50.

XX P-PSDB; ABG61940.

PT Detecting a prostate cancer-associated transcript in a cell in a

PT patient, useful for diagnosing prostate cancer (PC) or screening

PT modulators of PC, by determining if prostate cancer-associated genes

PT are expressed in a prostate tissue -

XX Claim 22; Page 421; 436pp; English.

XX The present invention relates to methods of detecting a prostate

CC cancer-associated transcript in a cell from a patient. The method

CC comprises contacting a biological sample from the patient with

CC prostate cancer-associated polynucleotides (designated PC genes) that

CC selectively hybridize to a sequence that is at least 80% identical

CC	to them. The prostate cancer-associated polynucleotide sequences
CC	are differentially expressed in prostate tumour tissue or in
CC	prostate cancer and are derived from the tissues of various
CC	organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC	The methods of the invention are useful for diagnosing and treating
CC	prostate cancer in mammals. The prostate cancer-associated genes are
CC	useful for diagnosing or treating prostate cancer, as well as for
CC	identifying modulators of prostate cancer or agents that inhibit
CC	prostate cancer. The nucleic acid sequences are particularly useful
CC	in gene therapy, as a vaccine or in antisense applications.
CC	ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC	sequences.
XX	
Sequence	3004 BP; 874 A; 679 C; 710 G; 741 T; 0 other;
Query Match	7.6%; Score 186.2; DB 24; Length 3004;
Best Local Similarity	55.0%; Pred. No. 7.8e-38;
Matches 443; Conservative	0; Mismatches 38; Indels 24; Gaps 3;
QY	1004 CTCAAGGGTCAAAATGATCTACTTACCTGAACAGATAGCATACTTTTTCATGTTTCACCA 1063
Db	1609 CTCAAAGGCCAAATGATCTACATGTTGTAATCCAGTCCAATCTGTGTTTGGGGTCACCC 1668
QY	1064 AGTGTGATGAACCTGGACGATTTTGCAAGAGAGAGGGCTGTATCTAACTGACATCCCTCTGT 1123
Db	1669 TGTGTGGACAGATTAGAAGATTTTACAGGACGAGGGCTCTACCTCTCAGACATCCCAATT 1728
QY	1124 CATGATGCCACGCGCGATCTGTCTTTTGGGAGACAATTTAGAGAGGAATACAAACTC 1183
Db	1729 CACAAATGCACTGAGGATGTGGTCTTAATAGGGGAACAGCCGAGCTCAAGATGGCGCTG 1788
QY	1184 ACCCAAGAACTGGAAATCCCTACTGCAGGCTACAGCTACAGTTCAGTTCAGAGCCCTGGAAGAT 1243
Db	1789 AAGAAGAGGCTGGGAAGCTGAAGCTTAAGCTTGAAGCAAGCCCAAGCCCTGGAGGAG 1848
QY	1244 GAAAGAAAAGACAGACACATTTGCTGTATTTCTGTCCTTCCCTCGGTCTGTTGCCAATGAG 1303
Db	1849 GAGAAGAAAAGACAGTAGACCTTCTGTGCTCCATATTTCCCTGTGAGGTGTGCTCAGCAG 1908
QY	1304 CTCGGGCACAAGCGTCCAGTGCCTGCCAAAGATATGACAACTGACCATCTCTTTTGTAGT 1363
Db	1909 CTGTGGCAAGGGCAAGTTGTGCAAGCCACAGAGTTCAGTATGTTCACCATGCTCTTCTCA 1968
QY	1364 GGCATTGTGGGCTTCAATGCTTTCTGTAGCAAGCATGCATCTGGAGAAGGAGCCATGAAG 1423
Db	1969 GACATCTTTGGGTTCACTGCCATCTGCTCCAGTGCTC-----ACCGCTGCAG 2016
QY	1424 ATGCTCAACCTCTCAACGACCTCTACACAGATTTGACACACTGACTGATTCGCCGAAA 1483
Db	2017 GTCATCACCATGCTCAATGCACGTACACTCGCTTCGACCAGCAGTGTGGAGAGCTGGA- 2075
QY	1484 AACCCATTTCTTTATAGGTGGAGACTGTTGGTGACAAGTATATGACAGTCACTGGTGTTA 1543
Db	2076 -----TGTCACAGGTGGAGACCAATTCGCAATGCTTATTTGTGTGGCTTGGGGGAATTA 2127
QY	1544 CCAGAGCCATGTCATTACCAATGCACGATCCATCTGCCACCTGGCCTTGGACATGATGGAA 1603
Db	2128 CACAAGAGAGTGTACTCATGCTGTTTCAGATAGCCGTGATGGCCCTGAAGATGATGGAG 2187
QY	1604 ATTGCTGGCCAGGTTTCAAGTA---GATGGTGAATCTGTTTCAGATAACAAATAGGGATACAC 1660
Db	2188 CTCTCTGATGAAGTTATGCTCTCCCATGGGAACCTTATCAAGATGCGAATTTGGACTGCAC 2247
QY	1661 ACTGGAGAGGTAGTTACAGGTGTTCATAGGACACGGGATGCGCTCGATCTGTTTTCGGG 1720
Db	2248 TCTGGATCAAGTTTTGCTGGCGTCGTGTGGAGTTTAAATGGCCCGTTACTGCTCTCTTTTGGGA 2307
QY	1721 AATACTGTCAACTCTCAACGCCCAACAGAAACACACAGGAGAAAAGGGGAAAATAAATGTG 1780
Db	2308 AACAAATGTCACTCTGGCTACAAATTTTGACTCTCTGCAAGTGTACCAAGAAAATCAATGTC 2367
QY	1781 TCTGAATATATACATACAGATGCTCTTA 1805

[illegible]

cc
pred

2368 AGCCCAACACTTACAGATTACTCA 2392

RESULT 15
ABK63662

ID ABK63662 standard; cDNA; 2335 BP.
AC ABK63662;
XX
XX 18-JUN-2002 (first entry)
XX
XX Rat sequence differentially expressed in response to a hepatotoxin #1569.
DE
XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW differential expression; centrilobular necrosis; steatosis.
XX
XX Rattus norvegicus.
OS
XX WO200210453-A2.
PN
XX 07-FEB-2002.
PD
XX 30-JUL-2001; 2001WO-US23872.
PF
XX 31-JUL-2000; 2000US-222040P.
PR
XX 02-NOV-2000; 2000US-244880P.
PR
XX 11-MAY-2001; 2001US-290029P.
PR
XX 15-MAY-2001; 2001US-290645P.
PR
XX 22-MAY-2001; 2001US-292336P.
PR
XX 06-JUN-2001; 2001US-295798P.
PR
XX 13-JUN-2001; 2001US-297457P.
PR
XX 19-JUN-2001; 2001US-298884P.
PR
XX 09-JUL-2001; 2001US-303459P.
XX
XX (GENE-) GENE LOGIC INC.
PA
XX Mendrick D, Porter MW, Johnson KR, Castile AL, Elashoff MR;
PI WPI; 2002-241625/29.
DR
XX Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or
PT cells exposed to the toxin and comparing these to gene expression in
PT unexposed tissues or cells -
XX
XX Claim 1; Seq ID No 1569; 239pp; English.
PS
XX The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic
CC effect of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of
CC the genes is indicative of at least one toxic effect or progression.
CC The method can also be used to identify an agent which modulates the
CC toxic response and predict cellular pathways that a compound modulates
CC in a cell. The methods utilise a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridises to a gene listed in the specification, a computer
CC system comprising a database containing information identifying the
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC set of genes comprising at least two genes listed in the specification,
CC and a user interface to view the information used to present information,
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cells exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell

CC sample that has been exposed to a compound or agent. Hepatotoxicity
CC is characterized by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
XX which is differentially expressed in response to a hepatotoxic agent.
SQ Sequence 2335 BP; 660 A; 535 C; 596 G; 544 T; 0 other;

Query Match 6.0%; Score 147.2; DB 24; Length 2335;
Best Local Similarity 51.0%; Pred. No. 9.9e-28;
Matches 517; Conservative 0; Mismatches 443; Indels 54; Gaps 5;

QY	725	TTCTGCAAGCTTTTCCTTTTCATATAATATTGACGGGACCTAGTGTCTACTCAGTGT	784
DB	771	TTCTGTGATGCTTTTCCTTCCACATGCTTTGATGAAGCACTAAGGGTCAAGCAAGCT	830
QY	785	GGCAATGCTATATACAGAGTTCTCCCGCAGCTCCAGCCTGGGAATTGCAGCCTTCTGTCT	844
DB	831	GGAGTGAATATTCAGAAGTATGTCCTCGAATCTTAAACCAGAGTTTGCACTAGATGAG	890
QY	845	GTCTTCTCGCTGGTTCGCTCATATTGATATTAGTTTCCATGGGATCCTTCTCACATC	904
DB	891	TATTTTTCATCATCCACCTCAAGTTACTTTCAACATCTCCAGCATCTGCAAGTTCATT	950
QY	905	AATACTGTTTTTGTATTGAGAGCAAGGAAGATTGTTGGATGTGGAGAAATTAGAAATGT	964
DB	951	RACAGTCAGTTTGTCTTGAGACAGAAAGAAATGATGCC-----AAAGC	997
QY	965	GAGGATGAAGTCACTGGGACTGAGATCAGTCTTACGTTCTCAAGGGTCAAAATGATCTAC	1024
DB	998	AAGGAAGAGCCAGCGGCTCTAAA-----CTCCGGGTGAGATGATCTGG	1043
QY	1025	TTACCTGAAGCAGATAGCATATTTTCTATGTTTCAACCAAGTGTCTGATGACCTGGACGAT	1084
DB	1044	ATGGAGTCTCTGAGGTGATGATCTCATGTGTTTCCCAACGTCCTGAGCCTGCAAGAG	1103
QY	1085	TTGACAAGGAGGCTGTATTAAGTGACATCTCTGATGATGCCACGCGCATCTT	1144
DB	1104	CTGGAAGAGAGCAAGATGATCTTTCTGATATCGTCCGACGACGACGAGGATCTC	1163
QY	1145	GTTCTTTTGGGAGACAATTTAGAGAGGAATACAACTCACCCAGAACTGGAAATCTCTC	1204
DB	1164	ATCTCTCTCAACAGCAGAGGCTGGCAGAGATGGAGTGTCTCTGCCAAGTGGAAAGAG	1223
QY	1205	ACTGACAGCTACAGCTCACGTTAAGAGCCCTGGGAAGATGAAAGAAAGAACACACACA	1264
DB	1224	AAGGAGGAGTTCGGTGTCTTTTCCAACTCACCTGGCCCATCGGAGAAGAAGACAGAGCC	1283
QY	1265	TTGCTGTATTCT	1324
DB	1284	TTGCTGTATGCCATGCTGCTGAACATGTGGCCACCACTCAAGGAGGGGAGAAGGTG	1343
QY	1325	CTGCGCAAAAGATATGACAATGTGACCATCTCTCTTTAGTGGCAATTTGGGGCTTCAATGCT	1384
DB	1344	GCTCGAGGAGAAATTTGAACATGTACAACTCTTTTTCAGCGATGTTGTGACATTTACCAAC	1403
QY	1385	TTCTGTAGCAAGCATGTCATCTGGAGAGGAGCCATGAGATCGTCAACCTCTCAACGAC	1444
DB	1404	ATCTGTGAGCGCTGTGAACCT-----ATCCAAATCTGAAACATGCTGAATCA	1451
QY	1445	CTCTACACAGATTGTGACACTGACTGATTCCTCCGGAACCCATTTGTTTATAGGTG	1504
DB	1452	ATGTAATCCAAAGTTTGACAGGTTTACCAGTGTCCATGA-----TGCTACAAAGTA	1502
QY	1505	GAGACTGTTGGTGACAAGTATATGACAGTGTGTTTACCAGAGCCATGCAATTCACCAT	1564
DB	1503	GAAACAATAGGGGATGCTTTACATGTTGGTGGGTGGAGTACCACTACCCGTTGAAAGCCAT	1562
QY	1565	GCACGATCCATCTGCCACCTGGCTTGGACATGA-----TGGAATGCTGGCCAGGTT	1618
DB	1563	GCTCAAGAGAGTCCCAATTTTCTCTGGGGATGAGAAATTTCTGCAAAAGAGTGTGAAT	1622
QY	1619	CAAGTAGATGTTGATCTCTTTCAGATAACAATAGGATACACACTGGAGGATGATTACA	1678

Search completed: July 1, 2003, 13:26:20
Job time : 357.866 secs

Db	1623	CCTGTCACTGGGGAACCTATCCAGATCAGATGGGAAATCCACACTGGACCAGCTCTTAGCA	1682
QY	1679	GGTGTATAGGACAGCGGATGCTCGATACTGCTTTTGGGAATACTGTCAAC	1732
DB	1683	GGTGTGTGGGAGACAAGATGCTCGTACTGCTGTTTGGTGACACACTGTAAAC	1736

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	2262.4	92.6	3068	11	EC026332	Homo sapi	BC026332
2	903.4	37.0	919	9	AL548673	AL548673	AL548673
3	840.6	34.4	1073	14	BQ421149	AGENCOURT	BQ421149
4	782.8	32.0	809	13	BI916583	603178512	BI916583
5	732.4	30.0	777	13	BI838266	603083189	BI838266
6	29.8	823	12	823	BGI199621	RST18913	BGI199621

ALIGNMENTS

[illegible]

REFERENCE	Mammalia, Eutheria, Plimates, Carnivora, Homio-
AUTHORS	1 (bases 1 to 3068)
TITLE	Strausberg, R.
JOURNAL	Direct Submission
	Submitted (02-APR-2002) National Institutes of Health, Mammalian
	Gene Collection (MGC), Cancer Genomics Office, National Cancer
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
	USA
REMARK	NIH-MGC Project URL: http://mgc.ncbi.nih.gov
COMMENT	Contact: MGC help desk
	Email: cgapbs-remail.nih.gov
	Tissue procurement: Miklos Palkovits, M.D., Ph.D.
	cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
	Toshiyuki and Piero Carninci (RIKEN)
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA sequencing by: Sequencing group at the Stanford Human Genome
	Center, Stanford University School of Medicine, Stanford, CA 94305
	Web site: http://www.shgc.stanford.edu

QY 1966 GAAGAGGATACAGACTAGTTCAGATTTCTCTTAACAGGTGCCAAGCCAGGAGCAGT 2025
Db 1884 GAAGAGGATACAGACTAGTTCAGATTTCTCTTAACAGGTGCCAAGCCAGGAGCAGT 1943
QY 2026 TCTTCCCTATGGATACAGATTTCTTTGTGCTCTGCTTACATACCCCAAGACTTTCTTCTA 2085
Db 1944 TCTTCCCTATGGATACAGATTTCTTTGTGCTCTGCTTACATACCCCAAGACTTTCTTCTA 2003
QY 2086 GATATATCTCTACATACCGTTATTAACCTTAGCTCTGCTTCTATATCTTTTAGGCT 2145
Db 2004 GATATATCTCTACATACCGTTATTAACCTTAGCTCTGCTTCTATATCTTTTAGGCT 2063
QY 2146 TTAGTATATATCTAAAGTTGGCTTTGATGTGGATGATGAGTTCATGCTCTCTAA 2205
Db 2064 TTAGTATATATCTAAAGTTGGCTTTGATGTGGATGATGAGTTCATGCTCTCTAA 2123
QY 2206 AATCTACTACAAGCATTACCTAACATGCTGATCTGCAAGTAGTAGGACCCCAATAATAT 2265
Db 2124 AATCTACTACAAGCATTACCTAACATGCTGATCTGCAAGTAGTAGGACCCCAATAATAT 2183
QY 2266 TTGTTGAATTTAGTTAAATGAACACTGAACAGTGTGGCCATGCTATATTTATATCATG 2325
Db 2184 TTGTTGAATTTAGTTAAATGAACACTGAACAGTGTGGCCATGCTATATTTATATCATG 2243
QY 2326 TTACCAAAATCTGTTAGTGTTCACATATATATGATATATATTTAAATGACTATAATG 2385
Db 2244 TTACCAAAATCTGTTAGTGTTCACATATATATGATATATATTTAAATGACTATAATG 2303
QY 2386 TAATAAAGTTTATATCATGTTGGTGTATATCATATATAGAAATCATTTCTTAAAGGAGT 2443
Db 2304 TAATAAAGTTTATATCATGTTGGTGTATATCATATATAGAAATCATTTCTTAAAGGAGT 2361

RESULT 2

AL548673 919 bp mRNA linear EST 16-FEB-2001
LOCUS AL548673 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0DI036YN13 5
DEFINITION Prime mRNA sequence.

ACCESSION AL548673

VERSION AL548673.1 GI:12883913

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 919)

Li, W.B.; Gruber, C.; Jesses, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

FEATURES

source

1. 919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI036YN13"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fullength@invitrogen.com
http://fulllength.invitrogen.com"

BASE COUNT

ORIGIN

252 a 204 c 207 g 254 t

2 others

RESULT 3

BQ421149

LOCUS

DEFINITION

AGENCOURT_7766609 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5015435

1073 bp mRNA linear

EST 23-MAY-2002

IMAGE:5015435

Query Match 37.0%; Score 903.4; DB 9; Length 919;
Best Local Similarity 99.5%; Pred. No. 1e-216;
Matches 915; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 17 CCGCCTCTGCTGGTCCCTTCGGCCGTACCTCTGCTGGGGGCTGCTCCCGGCTCC 76
Db 1 CCGCCTCTGCTGGTCCCTTCGGCCGTACCTCTGCTGGGGGCTGCTCCCGGCTCC 60
QY 77 GGTGCAGACACCATTGTAGGATTTGTGAATCACGCCCTGGAGTTGCTGGTATCGGCAAT 136
Db 61 GGTGCAGACACCATTGTAGGATTTGTGAATCACGCCCTGGAGTTGCTGGTATCGGCAAT 120
QY 137 TACGCCCCCGAGGTGTGGGAGACATCAAAAAGAGGCACAGTTAGATGAGAGAGCAG 196
Db 121 TACGCCCCCGAGGTGTGGGAGACATCAAAAAGAGGCACAGTTAGATGAGAGAGCAG 180
QY 197 TTTCTTGTCAAGATAATATATGATGACTCCAAAATCTTATGATTTGGTGTCTGCTGCAAGC 256
Db 181 TTTCTTGTCAAGATAATATATGATGACTCCAAAATCTTATGATTTGGTGTCTGCTGCAAGC 240
QY 257 AAAGTCTCTCAATCTCAATGCTGGAGAAATCTCCAAAATGTTTGGGAGAGATGTTTTTCGTC 316
Db 241 AAAGTCTCTCAATCTCAATGCTGGAGAAATCTCCAAAATGTTTGGGAGAGATGTTTTTCGTC 300
QY 317 TTTTGGCAAGAAATCTGTTATGATCAATCTTGGTGTCTGCTGGGCTCTAATGTCAGAA 376
Db 301 TTTTGGCAAGAAATCTGTTATGATCAATCTTGGTGTCTGCTGGGCTCTAATGTCAGAA 360
QY 377 TTTCTACAGAACCTTGTATGCTCTGCAGCACCACCTTGTACCATCTACCCAGGAATGGT 436
Db 361 TTTCTACAGAACCTTGTATGCTCTGCAGCACCACCTTGTACCATCTACCCAGGAATGGT 420
QY 437 GCACCTTCTTTAGGTGCACATGTCAGAAAAGGCAAGGACTCATTTTGCACCTACTAC 496
Db 421 GCACCTTCTTTAGGTGCACATGTCAGAAAAGGCAAGGACTCATTTTGCACCTACTAC 480
QY 497 TCAGAGAGAGAGGACTTCAGGATATTTGTCATTTGGAATCATCAAAAGAGTGGCACAACAA 556
Db 481 TCAGAGAGAGAGGACTTCAGGATATTTGTCATTTGGAATCATCAAAAGAGTGGCACAACAA 540
QY 557 ATCCATGGCACTGAAATAGACATGAAGGTTATTTCAGCAAAAGAAATGAAGAATGTGATCAT 616
Db 541 ATCCATGGCACTGAAATAGACATGAAGGTTATTTCAGCAAAAGAAATGAAGAATGTGATCAT 600
QY 617 ACTCAATTTTAAATGAAGAAAAGAGTCAAAAGAGAGGATTTTATGAAGATCTTGAC 676
Db 601 ACTCAATTTTAAATGAAGAAAAGAGTCAAAAGAGAGGNTTTTATGAAGATCTTGAC 660
QY 677 AGATTTGAAGAAATGTTACCCAGGAATCAGCATCAGCCCATATACATTTCTGCAAGCT 736
Db 661 AGATTTGAAGAAATGTTACCCAGGAATCAGCATCAGCCCATATACATTTCTGCAAGCT 720
QY 737 TTTCTTTTTCATATAATATTTGACCGGACCTAGTGGTCACTAGTGGCAATGCTATA 796
Db 721 TTTCTTTTTCATATAATATTTGACCGGACCTAGTGGTCACTAGTGGCAATGCTATA 780
QY 797 TACAGAGTTCTCCCCAGCTCCAGCTGGGAATTCAGACCTTCTGCTGTCTTCTTCGCTG 856
Db 781 TACAGAGTTCTCCCCAGCTCCAGCTGGGAATTCAGACCTTCTGCTGTCTTCTTCGCTG 840
QY 857 GTTCGCTCTCATATATGATTTAGTTTCCATGGATCCCTTTCTCATCAATACATCTTTTT 916
Db 841 GTTCGCTCTCATATATGATTTAGTTT-CATGGATCCCTTTCTCATCAATACATCTTTTT 899
QY 917 GTATTGAGAAGCAAGGAGG 936
Db 900 GTATTGAGAAGCAAGGAGG 919

```

5', mRNA sequence.
B0421149
VERSION B0421149.1 GI:21116464
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13211 row: 1 column: 04
High quality sequence stop: 617.
Location/Qualifiers
1. .1073
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6015435"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: testis; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 300 a 236 c 262 g 275 t
ORIGIN
Query Match 34.4%; Score 840.6; DB 14; Length 1073;
Best Local Similarity 93.5%; Pred. No. 6.9e-201;
Matches 932; Conservative 0; Mismatches 59; Indels 6; Gaps 5;
QY 559 CCATGGCACTGAATAGACATGAAGTTATTTCAGCAAGAAATGAAGATGTGATCATAC 618
DB 10 CGATGGCACTGAATAGACATGAAGTTATTTCAGCAAGAAATGAAGATGTGATCATAC 69
QY 619 TCAATTTTAAATGAAGAAAGAGTCAGCAAGAGGATTTTATGAGATCTTGACAG 678
DB 70 TCAATTTTAAATGAAGAAAGAGTCAGCAAGAGGATTTTATGAGATCTTGACAG 129
QY 679 ATTTGAAGAAATGGTACCAGCAATCAGCATCAGCCCATATACATCTGCAAAAGCTTT 738
DB 130 ATTTGAAGAAATGGTACCAGCAATCAGCATCAGCCCATATACATCTGCAAAAGCTTT 189
QY 739 TCGTTTTCATATATATTTGACCGGACCTAGTGTGCTAGTGTGGAATGCTATATA 798
DB 190 TCGTTTTCATATATATTTGACCGGACCTAGTGTGCTAGTGTGGAATGCTATATA 249
QY 799 CAGAGTTCTCCCGCCAGCTCCAGCTGGGAATGACGCTTCGTCGTCCTCGCTGT 858
DB 250 CAGAGTTCTCCCGCCAGCTCCAGCTGGGAATGACGCTTCGTCGTCCTCGCTGT 309
QY 859 TCGTCTCATATATATATTTGATTTCCATGGGATCCCTTTTCACATCAATACATCTTTGT 918
DB 310 TCGTCTCATATATATATTTGATTTCCATGGGATCCCTTTTCACATCAATACATCTTTGT 369
QY 919 ATTGAGACAGGAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 978
DB 370 ATTGAGACAGGAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 429
QY 979 TGGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038

```

```

Db 430 TGGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489
QY 1039 TAGCATACATTTTCTATGTTTACCAAGTGTCTATGAACCTGGAGGATTTTACCAAGGAGG 1098
DB 490 TAGCATACATTTTCTATGTTTACCAAGTGTCTATGAACCTGGAGGATTTTACCAAGGAGG 549
QY 1099 GCTGTATCTAAGTGAATCCCTCTGTCATGATCCGCGCGATCTTGTGTTTGTGGGAGA 1158
DB 550 GCTGTATCTAAGTGAATCCCTCTGTCATGATCCGCGCGATCTTGTGTTTGTGGGAGA 609
QY 1159 ACAATTTAGAGAGGATACAAACTCA-CCCAAGAACTGAAATCTCTACTGACAGGCTAC 1217
DB 610 ACAATTTAGAGAGGATACAAACTCA-CCCAAGAACTGAAATCTCTACTGACAGGCTAC 669
QY 1218 AGCTCAGCTTAAGAGCCCTGGGAAGATGAAAGAAAGACACACATCTGTTGTTTCTG 1277
DB 670 AGCTCAGCTTAAGAGCCCTGGGAAGATGAAAGAAAGACAGGACCCCTTCTGTTATCTG 729
QY 1278 TCCCTTCCCTCCGCTCTGTTGCCAATGAGCTGCGGCACAAAGCTCCAGTGCCTGCCAAAGAT 1337
DB 730 TCCCTTCCCTCCGCTCTGTTGCCAATGAGCTGCGGCACAAAGCTCCAGTGCCTGCCAAAGAA 789
QY 1338 ATGACAATGTG-ACCATCTCTTTAGTGCATTTGGGCTTCAATGCTTTCTGTGACCAAG 1396
DB 790 GTGACAATGTGAACCATCTCTTAGAGGGCATTTGGGGCTTCAAGGCTTTCTGTACAAAG 849
QY 1397 CATGATCTGGAGAGGAGGACCATGAAGATCGT-CAACCTCTCTCAACGACCTCTACACCAAG 1455
DB 850 CAGGCATTTGGAGAGGAGGACCATGAATCGTCCAAACCTCTCAAGGAATCTCTACCAAG 909
QY 1456 ATTT-GACACATGATGATTTCCCGGAAAAACCCATTTG--TTTATAAGTGGAGACTGT 1512
DB 910 ATTTGGCAACACCGCATGAATTTCCGGGAGAAAACCTTTTGGTTTATAGGGGGGAATGCT 969
QY 1513 TCGTGACAAATATATGACAGTGAAGTGTGTTTACCAAG 1549
DB 970 TGTGCCCAAGTATTTGAACCGGGGTGGTGTATACCCAG 1006

```

```

RESULT 4
B1916583
LOCUS B03178512F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5242779 5',
DEFINITION mRNA sequence.
ACCESSION B1916583
VERSION B1916583.1 GI:16199460
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Inocyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11612 row: g column: 04
High quality sequence stop: 804.
Location/Qualifiers
1. .809
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5242779"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"

```



```

Db 300 GATTGTGGATGGAGAAATAGAAATGTAGAGATGAAGTGAAGTGGAGTGAAGTGAAGT 359
QY 996 GCTTACGCTCTCAAGGCTCAAAATGATCTACTTACCTGAAGCAGATAGATAGTCTTCTAT 1055
Db 360 GCTTACGCTCTCAAGGCTCAAAATGATCTACTTACCTGAAGCAGATAGATAGTCTTCTAT 419
QY 1056 GTTACCAAGTGTATGAACCTGGAGCATTTGACAAAGGAGAGGGCTGTATCTAAGTGACA 1115
Db 420 GTTACCAAGTGTATGAACCTGGAGCATTTGACAAAGGAGAGGGCTGTATCTAAGTGACA 479
QY 1116 TCCTCTGATGATGATGCCAGCGGATCTTGTCTTTTGGAGAGCAATTTAGAGGAAT 1175
Db 480 TCCTCTGATGATGATGCCAGCGGATCTTGTCTTTTGGAGAGCAATTTAGAGGAAT 539
QY 1176 ACAAACTCAACCAAGAACTGGAAATCCTCAGTACAGAGGCTACAGCTCAGCTTAAAGAGCCC 1235
Db 540 ACAAACTCAACCAAGAACTGGAAATCCTCAGTACAGAGGCTACAGCTCAGCTTAAAGAGCCC 599
QY 1236 TGAAGATGAAAGAAAAGACAGACATTTGCTGTATTTCTTCTTCTTCTTCTTCTTCTTCTT 1295
Db 600 TGAAGATGACAAAGAAACAGACACATTTGCTGTATTTCTTCTTCTTCTTCTTCTTCTTCTT 659
QY 1296 CCAATGAGCTGGGCAACAGGCTCCAGTGCCTGCCAAAGATATGACATGTGACCATCC 1355
Db 660 CCAATGAGCTGGGCAACAGGCTCCAGTGCCTGCCAAAGATATGACATGTGACCATCC 719
QY 1356 TCTTTAGTGGCATTTGGGCTTCAATGCTTCTGTAGCAAGCATGCTGGAGAGG 1413
Db 720 TCTTTAGTGGCATTTGGGCTTCAATGCTTCTGTAGCAAGCATGCTGGAGAGG 777

```

RESULT 6

```

BG199621/c 823 bp mRNA linear EST 21-APR-2001
LOCUS RST18913 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION BG199621
ACCESSION BG199621.1 GI:13721308
VERSION 1
KEYWORDS EST.
SOURCE human.

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 823)
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith,
E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,
J., Danzig, J., and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
2127151
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scaine@atersys.com
High quality sequence stop: 348.
Location/Qualifiers
1..823
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
249 a 165 c 146 g 263 t

```

REFERENCE

```

AUTHORS

```

TITLE

```

JOURNAL
MEDLINE
COMMENT

```

FEATURES

```

source

```

```

Query Match 29.8%; Score 728; DB 12; Length 823;
Best Local Similarity 96.4%; Pred. No. 1.6e-172;
Matches 707; Conservative 0; Mismatches 25; Indels 4; Gaps 4;
QY 1586 GCCTTGGACATGATGGAATTCCTGGCCAGGTTCAAGTAGATGGTGAATCTGTTCAGATA 1645
Db 821 GCCTTGGACATAAGTTAAAT-CTGCTCAGGTTCAAGTCGATGGTAAATCTGTTCAGATA 763
QY 1646 ACAATAGGATACACACTGGAGAGGTAGTACAGGTGTATAGGACAGGGATGCCCTCA 1705
Db 762 ACTATAGGATACACACTGGAGAGGTAGTACAGGTGTATAGGACAGGGATGCCCTCA 703
QY 1706 TACTGTCTTTTGGGAATCTGTCAACCTCACAGCCGACAGAAACACAGAGGAGAAAG 1765
Db 702 TAATGTCTTTTGGGAATCTGTCAACCTCACAGCCGATCAGAAACACAGAGGAGAAAG 643
QY 1766 GGAATAAATGTGTCTGAATATACATACAGATGTCTTATGTCTCCAGAAATTCAGAT 1825
Db 642 GGAATAAATCAATGTCTCTGAATATACATACAGATGTCTTATGTCTCCAGAAATTCAGAT 583
QY 1826 CCACAACTCCACTTGGAGCACAGAGCCAGTCCATGAAGGCAAGGCAAGGCAAGGCAAGG 1885
Db 582 CCACAACTTACACTTGGAGCACAGAGCCAGTCCATGAAGGCAAGGCAAGGCAAGGCAAGG 523
QY 1886 CAAGTTTGGTTTCTATCCAGAAAAATACAGGAACAGAGAAACAAAGCAGGATGATGAC 1945
Db 522 CAAGTTTGGTTTCTATCCAGAAAAATACAGGAACAGAGGAAACAAAGCAGGATGATGAC 463
QY 1946 TGAATCTTGATATGGGTGAAGAGGATACAGACTAGTTCACAGTTCCTTCCTTCAACAC 2005
Db 462 TGAATCTTGATATGGGTGAAGAGGATACAGACTAGTTCCTTCCTTCAACAC 403
QY 2006 GTGCCAAGCCAGGAGCAGTCTTCCCTATGATACAGATTTCTTCTTCTTCTTCTTCTTCT 2065
Db 402 GTGCCAAGCCAGGAGCAGTCTTCCCTATGATACAGATTTCTTCTTCTTCTTCTTCTTCT 343
QY 2066 TACCCCAAGACTTCTCTAGATATATCTCTCACTATCCCTATTCCTTATTCCTTATTCCT 2125
Db 342 TACCCCAAGACTTCTCTAGATATATCTCTCACTATCCCTTATTCCTTATTCCTTATTCCT 283
QY 2126 TTTCTATTAATCTTTAGGCTTTAGTATATATCTTAAAGTTTGGCTTTTGGTGGATGAT 2185
Db 282 TATCTATTAATCTTTAGGCTTTAGTATATATCTTAAAGTTTGGCTTTTGGTGGATAAT 223
QY 2186 GTGAGCTTCATCTGTCTTAAATCTACTACAGCATTACCTAACATGGTGTATCTGCAACT 2245
Db 222 GTGAGCTTCATCTGTCTTAAATCTACTACAGCATTACCTAACATGGTGTATCTGCAACT 163
QY 2246 AGTAGGACCCCAATAAATTTTGAATTTAGTTAAATGAAACTGAACAGTGTGTG-GC 2304
Db 162 AGTAGGACCCCAATAAATATATGTGCAATTTAGTTAAATGAAACTGAACAGTGTGTGTC 103
QY 2305 CATGTGTATATTTATATCATGTTTACCAG-ATCTGTTTGTGTGTCCA-CATATATGATA 2362
Db 102 CATGTGTATATTTATATCATGTTTACCAGATCTGTTTGTGTGTCCAACCATATATGATA 43
QY 2363 TGTATATTTTAAATGACTATAATGTAATAAAGTTTAT 2398
Db 42 TGTATATTTTAAATGACTATAATGTAATAAAGTTGAT 7

```

RESULT 7

```

BI666927 810 bp mRNA linear EST 12-SEP-2001
LOCUS 603291676f1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:531146 5',
DEFINITION mRNA sequence.
ACCESSION BI666927
VERSION BI666927.1 GI:15581160
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 810)
 NIH-MGC <http://mhc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11788 row: g column: 19
 High quality sequence stop: 770.
 Location/Qualifiers
 1. 810
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIMH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 244 a 168 c 175 g 223 t

BASE COUNT
 ORIGIN

Query Match 27.9%; Score 680.8; DB 13; Length 810;
 Best Local Similarity 97.4%; Pred. No. 1.3e-160;
 Matches 756; Conservative 0; Mismatches 12; Indels 8; Gaps 6;

QY 166 AAAAGAGCACAGTTAGATGAAGAGGACAGTTCTTCTGCAGATAATATATGATGACTC 225
 DB 166 AAAAGAGCACAGTTAGATGAAGAGGACAGTTCTTCTGCAGATAATATATGATGACTC 98
 QY 226 CAAACTATGATTTGGTGTGCTGCAAGCAAGTCTCAATCTCAATGCTGGAGAAAT 285
 DB 99 CAAACTATGATTTGGTGTGCTGCAAGCAAGTCTCAATCTCAATGCTGGAGAAAT 158
 QY 286 CCTCCAAATGTTGGGAAGATGTTTTCGTCTTTTGGCAAGAACTGTTATGATACAAAT 345
 DB 159 CCTCCAAATG-TTGGGAAGATGTTTTCGTCTTTTGGCAAGAACTGTTATGATACAAAT 216
 QY 346 CTTCGTGCTCTGGGCTCTAATGTGAGAGATTTCTACAGAACCTTGATGCTCTCCAGCA 405
 DB 217 CTTCGTGCTCTGGGCTCTAATGTGAGAGATTTCTACAGAACCTTGATGCTCTCCAGCA 276
 QY 406 CCACCTTGCTACCATCTACCCAGGAATCGGTCACCTCTCTTTAGGTCACTGATGACAGA 465
 DB 277 CCACCTTGCTACCATCTACCCAGGAATCGGTCACCTCTCTTTAGGTCACTGATGACAGA 336
 QY 466 RAAGGCAAGGACTCATTTTGCATCTACTCTACAGAGAGAGGACCTTCAGGATATGCT 525
 DB 337 RAAGGCAAGGACTCATTTTGCATCTACTCTACAGAGAGAGGACCTTCAGGATATGCT 396
 QY 526 CATTGGAATCATCAAAACAGTGGCACAAATCATGCACTGGAATAGACATCAAGGT 585
 DB 397 CATTGGAATCATCAAAACAGTGGCACAAATCATGCACTGGAATAGACATCAAGGT 456
 QY 586 TATTCAGCAAGAAATGAAGATGTGATCATCTACTCAATTTTAAATGAAGAAAAAGATGC 645
 DB 457 TATTCAGCAAGAAATGAAGATGTGATCATCTACTCAATTTTAAATGAAGAAAAAGATGC 516

QY 646 AAAAGAAGAGGATTTTATGAAGATCTTGACAGATTTGAAGAAATGTACCCAGGAATC 705
 DB 517 AAAAGAAGAGGATTTTATGAAGATCTTGACAGATTTGAAGAAATGTACCCAGGAATC 576
 QY 706 ACGCATGACCCATACATATCTGCAAGAGCTTTTCTCTTTCATATAATATTTGACCCGGA 765
 DB 577 ACGCATGACCCATACATATCTGCAAGAGCTTTTCTCTTTCATATAATATTTGACCCGGA 636
 QY 766 CCTAGTGTCTACTCAGTGTGGCAATGCTATATACAGAGTTCTCCGCCAGCTCCAGCCCTGG 825
 DB 637 CCTAGTGTCTACTCAGTGTGGCAATGCTATATACAGAGTTCTCCGCCAGCTCCAGCCCTGG 696
 QY 826 GAATTCGAGCC-TTCTGTCTGTCTCTCG--CTGTTTCTGCTCTCATATTTGA-TATTAGTT 881
 DB 697 GAATTCGAGCCTTTCTGTCTGTCTCTCTGCTGCTGCTGCTCTCATATTTGATTAAGATT 756
 QY 882 TCCATGGGATCCTTCTCAGATCAATCACTGTTTGTATTGAGAGCAAGGAAGA 937
 DB 757 TCCATGGGATCCTTATATACATCACTG--GTTGGATTGAGAGCAAGGAAGA 810

RESULT 8
 BI553448
 LOCUS 603193201F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5264160 5',
 DEFINITION mRNA sequence.
 ACCESSION BI553448
 VERSION BI553448.1 GI:15440760
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 717)
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11666 row: b column: 01
 High quality sequence stop: 712.
 Location/Qualifiers
 1. 717
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_95"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.5 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIMH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 207 a 157 c 172 g 181 t

BASE COUNT
 ORIGIN

Query Match 27.9%; Score 680.4; DB 13; Length 717;
 Best Local Similarity 98.3%; Pred. No. 1.6e-160;
 Matches 698; Conservative 0; Mismatches 11; Indels 1; Gaps 1;


```

Db      720 AGGCTCCA-TTTCCTCCTAACAGTTCACAGCCGAGCACCA 759

RESULT 10
BO948975
LOCUS   AGENCOURT_8879414 NCI_CGAP_Co24 Mus musculus cDNA clone
DEFINITION IMAGE:6475913 5', mRNA sequence.
ACCESSION BO948975
VERSION   BO948975.1 GI:22364453
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 939)
           NIH-MGC http://mgc.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: The Cepko Laboratory
           cDNA Library Preparation: Life Technologies, Inc.
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L14M14015 row: g column: 18
           High quality sequence stop: 658.
FEATURES
source
1..939
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6475913"
/clone.lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 271 a 222 c 232 g 210 t 4 others
ORIGIN

Query Match 26.9%; Score 657.6; DB 14; Length 939;
Best Local Similarity 87.7%; Pred. NO. 9.1e-155;
Matches 739; Conservative 0; Mismatches 102; Indels 2; Gaps 2;

QY 1017 TGATCTACTTACCTGAAGCAGATAGCATACATCTTTTCTATGTTCACCAAGTGCATGAACC 1076
Db 12 TGATCTACTTACCAAGCAGATAGCATCCTCTTCTCTGTTCACCAAGTGCATGAACC 71
QY 1077 TGGAGGATTTGACAGGAGAGGGCTGTATCTAAGTGACATCCCTCTGCATGATGCCAGCC 1136
Db 72 TGGAGGACCTACAGAGAGAGGCCGTGTATCTGAGTGACATCCCTCTCCACGATGCTACCC 131
QY 1137 GCGATCTTGTCTTTTGGGAGAACAAATTAGAGAGGAATAACAACATCCACCAAGAACATGG 1196
Db 132 GAGACCTGGTCTTTTGGGAGAACAGTTCCGGGAGGAGTACAAACTGACACAGAGCTGG 191
QY 1197 AATCCTCAGTCAGAGGCTACAGCTCAGCTTAAAGAGCCCTGGAAGATGAAGAAAGAAAAGA 1256
Db 192 AATCCTCAGCAGAGGCTGACGCTCACACTGAGAGCCCTTGGAGGATGAGAAGAAAAGA 251
QY 1257 CAGACACATGCTGTATCTGTCCCTTCCCTGCTGTGCGCAATGAGCTGGCGCACAGC 1316
Db 252 CAGACACATGCTGTATCTGTCCCTTCCCTGCTGTGCGCAATGAGCTGGAGACACAAGC 311
QY 1317 GTCCAGTGCCTGCCAAAAGATGACAAATGTGACCATCTCTTTAGTGGCATTTGGGCT 1376
Db 312 GCCCAGTGCCTGCCAAAAGATGACAAATGTGACCATCTCTTTCAGCGGCATTGGGCT 371

QY 1377 TCAATGCTTTCTGTAGCAAGCATGCAATCTGGAGAGGAGGCCATGAAGATCGTCAACCTCC 1436
Db 372 TCAATGCTTTCTGTAGCAAGCATGCAATCTGGAGAGGAGGCCATGAAGATCGTCAATCTCC 431
QY 1437 TCAAGGACCTCTACACACAGATTTGACACACACTGACTGATTCGCGGAAAAACCAATTTGTT 1496
Db 432 TCAAGGATCTCTACACCCGATTTGACACACACTGACTGATTCAGGAAAAAACCAATTTGTT 491
QY 1497 ATAAGGTGGAGACTTGTGGTGACAAGTATATGACACAGTGAAGTGTGTTTACAGAGCCATGCA 1556
Db 492 ACAAGGTGGAACAGATTGTTGTCACAAGTATATGACACAGTGAAGTGTGTTTACAGAACCTGTA 551
QY 1557 TTCACCATGCAGCATCCATCTGCCACCTGGCCTTGGACATGATGAAGATTTGTCGCCAGG 1616
Db 552 TCCACCATGCAGCGTCCATTTGCCACCTGGCTTTAGACATGATGAAGATTTGTCGCCAGG 611
QY 1617 TTCAAGTAGAGTGGTGAATCTGTTCAAGATAACAATAGGACATACACACTGGAGAGGTAGTGA 1676
Db 612 TTCAAGTAGATGGTGAATCTGTTCAAGATAACAATCGGATCCATACGGGGAGGTGGTGA 671
QY 1677 CAGGTGTCATAGGACAGCGGATGCTCGATCTGCTTTTGGGAAATACTGTCAACCTCA 1736
Db 672 CAGGTGTCATAGGACAGCGGATGCTCGATCTGCTTTTGGG-ATACCGTCAACCTCA 730
QY 1737 CAAGCCGACAGAAACACACAGGAGAAAGGAAAAATAAATGTGTCTGAATATACATAC- 1795
Db 731 CAAGCAGGACAGAAACACACAGGAGAAAGGAAAAATAAATGTGTCTGAATATACATACA 790
QY 1796 AGATGCTCTATGCTCTCCAGAAAAATTCAGATCCACAATTCCTACTTGGACACAGAGGCCCA 1855
Db 791 AGGTGCTCTATGCTCTCCAGAAAAATTCAGATCCACTGTCTTCATTTGAGGACACAGAACCCA 850
QY 1856 GTG 1858
Db 851 ANG 853

RESULT 11
BF036708 749 bp mRNA linear EST 20-OCT-2000
LOCUS 601459903F1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3863401 5',
DEFINITION mRNA sequence.
ACCESSION BF036708
VERSION BF036708.1 GI:10744768
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 749)
           NIH-MGC http://mgc.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: DCTD/DP
           cDNA Library Preparation: Life Technologies, Inc.
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L14M9603 row: e column: 02
           High quality sequence stop: 684.
FEATURES
source
1..749
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3863401"
/clone.lib="NIH_MGC_66"
/tissue.type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: NotI;

```

Site 2: Sali; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.8 kb. Library constructed by Life
Technologies.

BASE COUNT 221 a 159 c 180 g 189 t

Query Match 26.3%; Score 641.8; DB 12; Length 749;
Best Local Similarity 97.8%; Pred. No. 8.6e-151;
Matches 682; Conservative 0; Mismatches 12; Indels 3; Gaps 3;
QY 1040 AGCATACTTTTCTATGTTCCACCAAGTGTCAACCTGGACGATTGACAGGAGGG 1099
Db 1 AGCATACTTTTCTATGTTCCACCAAGTGTCAACCTGGACGATTGACAGGAGGG 60
QY 1100 CTGTATCTAAGTGACATCCTCTGTCATGATGCCGCGGATCTTCTTTTGGGAGAA 1159
Db 61 CTGTATCTAAGTGACATCCTCTGTCATGATGCCGCGGATCTTCTTTTGGGAGAA 120
QY 1160 CAATTTAGAGAGATACAACTACCCCAAGACTGGAATCCTCACTGACAGGCTACAG 1219
Db 121 CAATTTAGAGAGATACAACTACCCCAAGACTGGAATCCTCACTGACAGGCTACAG 180
QY 1220 CTCACCTTAAGCCCTGCAAGTGAAGTGAAGAAAGACAGACATTCGTATCTGTC 1279
Db 181 CTCACCTTAAGCCCTGCAAGTGAAGTGAAGAAAGACAGACATTCGTATCTGTC 240
QY 1280 CTTCTCCCTCTGTTGCCAATGAGCTGCGGCACCAAGCTCCAGTGCCTGCCAAAGATAT 1339
Db 241 CTTCTCCCTCTGTTGCCAATGAGCTGCGGCACCAAGCTCCAGTGCCTGCCAAAGATAT 300
QY 1340 GACAATGTGACCATCCTTTTAGTGGCATGTGGGCTTCAATGCTTCTGTAGCAAGCAT 1399
Db 301 GACAATGTGACCATCCTTTTAGTGGCATGTGGGCTTCAATGCTTCTGTAGCAAGCAT 360
QY 1400 GCATCTGGAGAGGACCATGATGATGCTCAACCTCTCAACGACCTCTACACCATTT 1459
Db 361 GCATCTGGAGAGGACCATGATGATGCTCAACCTCTCAACGACCTCTACACCATTT 420
QY 1460 GACACACTGACTGATTTCCCGGAAAAACCCATTTGTTTATAGGTGGAGACTGTGGTGAC 1519
Db 421 GACACACTGACTGATTTCCCGGAAAAACCCCA-TTGTATTATAGGTGGAGACTGTGGTGAC 479
QY 1520 AGTATATCAGATGAGTGGTTTACCAGAGCCATGATTCACCATGACAGTCCATCTGC 1579
Db 480 AGTATATCAGATGAGTGGTTTACCAGAGCCATGATTCACCATGACAGTCCATCTGC 539
QY 1580 CACCTGGCCTTGACATGATGAAATGTCTGGCCAGGTTCAAGTAGATGTGAATCTGTT 1639
Db 540 CACCTGGCCTTGACATGATGAAAT-TGCTGGCCAGGTTCAAGTAGATGTGAATCTGTT 598
QY 1640 CAGATAACATAGGGATACACACTGGAGAGGTAGTTACAGGTGTCTATAGGACAGCGGATG 1699
Db 599 CAGATAACATAGGGATACACACTGGAGAGGTAGTTACAGGTGTCTATAGGACAG-GGATG 657
QY 1700 CCTCGATATGCTCTTTTGGGATACTGTCAACCTCA 1736
Db 658 CTCGATAATGTTCTTTTGGGATACTGTAACTTAA 694

RESULT 12
BG706547 721 bp mRNA linear EST 07-MAY-2001
LOCUS 602672916f1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795634 5',
DEFINITION mRNA sequence.
ACCESSION BG706547
VERSION BG706547.1 GI:13982001
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 721)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.

TITLE
JOURNAL
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10678 row: p column: 03
High quality sequence stop: 721.

FEATURES
Source

Location/Qualifiers
1..721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4795634"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: Sali-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 213 a 156 c 161 g 191 t
ORIGIN

Query Match 25.8%; Score 631; DB 12; Length 721;
Best Local Similarity 100.0%; Pred. No. 4.5e-148;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 AAAAGAGGCACAGTTAGATGAAGAAGGACAGTCTTGTGTGAGAATAATATATGATGACATC 225
Db 91 AAAAGAGGCACAGTTAGATGAAGAAGGACAGTCTTGTGTGAGAATAATATATGATGACATC 150
QY 226 CAAACTTATGATTTGGTGTGCTGCTGAAGCAAGTCTCAATCTCAATCTGCTGGAGAAAT 285
Db 151 CAAACTTATGATTTGGTGTGCTGCTGAAGCAAGTCTCAATCTCAATCTGCTGGAGAAAT 210
QY 286 CTTCCAAATGTTGGGAAGATGTTTTCCTCTTTTCCCAAGATCTGTTATGATACAT 345
Db 211 CTTCCAAATGTTGGGAAGATGTTTTCCTCTTTTCCCAAGATCTGTTATGATACAT 270
QY 346 CTTGGTGTCTCTGGGCTCTAATGTGAGAGAATTTCTACAGAACCTTTGATGCTCTGCACGA 405
Db 271 CTTGGTGTCTCTGGGCTCTAATGTGAGAGAATTTCTACAGAACCTTTGATGCTCTGCACGA 330
QY 406 CCACCTTGTACATCTACCCAGGAATGGTGCACCTTCCCTTTAGTGTGACATGATGACGA 465
Db 331 CCACCTTGTACATCTACCCAGGAATGGTGCACCTTCCCTTTAGTGTGACATGATGACGA 390
QY 466 AAAGGGCAAGGACCTCATTTTTCACCTACTACTACTACTACTACTACTACTACTACTACT 525
Db 391 AAAGGGCAAGGACCTCATTTTTCACCTACTACTACTACTACTACTACTACTACTACTACT 450
QY 526 CATTGGAATCATCAAAACAGTGGCACAACAAATCCATGCTGCTGCTGCTGCTGCTGCTGCT 585
Db 451 CATTGGAATCATCAAAACAGTGGCACAACAAATCCATGCTGCTGCTGCTGCTGCTGCTGCT 510
QY 586 TATTGAGCAAAAGATGAAGATGTGATCATCTACTACTACTACTACTACTACTACTACTACT 645
Db 511 TATTGAGCAAAAGATGAAGATGTGATCATCTACTACTACTACTACTACTACTACTACTACT 570
QY 646 AAAAGCAAGAGGATTTTATGAAGATCTTGCACAGATTTTGAAGAAAATGTTACCCAGGAATC 705
Db 571 AAAAGCAAGAGGATTTTATGAAGATCTTGCACAGATTTTGAAGAAAATGTTACCCAGGAATC 630

QY	706	ACGCATCAGCCCATATACATCTCTGCAAAAGCTTTTCCTTTTCATATAATATTGACCGGGA	760
Db	631	ACGCATCAGCCCATATACATCTCTGCAAAAGCTTTTCCTTTTCATATAATATTGACCGGGA	690
QY	766	CCTAGTGGTCACTCAGTGTGGCAATGCTATA	796
Db	691	CCTAGTGGTCACTCAGTGTGGCAATGCTATA	721
RESULT 13			
LOCUS	BG434435		
DEFINITION	602506454F1 NIH_MGC_79 Homo sapiens CDNA clone IMAGE:4603908 5', mRNA sequence.	808 bp	linear EST 14-MAR-2000
ACCESSION	BG434435		
VERSION	1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 808)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs@email.nih.gov		
	Tissue Procurement: CLONTECH Laboratories, Inc.		
	cDNA Library Preparation: CLONTECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Plate: L1CMI338 row: k column: 13		
	High quality sequence stop: 698.		
FEATURES			
source	1..808		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4603908"		
	/clone_lib="NIH_MGC_79"		
	/lab_host="DH10B (T1 phage-resistant)"		
	/notes="Organ: placenta; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggccatcatggcc) 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGCGCCGAGCGGCATG-3' (30)BN-3' (where B = A, C, G, or N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."		
BASE COUNT	242 a 179 c 181 g 206 t		
ORIGIN			
Query Match	25.4%;	Score 621;	DB 12: Length 808;
Best Local Similarity	94.5%;	Pred. No. 1.5e-145;	
Matches	763;	Conservative 0;	Mismatches 30; Indels 14; Gaps 11
QY	1370	GTGGGCTTCATGCTTTCTGTGAGCAAGCATCGATCGAGGAGGCCATGAAGATCGTC	1420
Db	1	GTGGGCTTCATGCTTTCTGTGAGCAAGCATCGATCGAGGAGGCCATGAAGATCGTC	60
QY	1430	AACCTCTCAACGACCTCTACACAGATTTTGACACACTGACTGATTCCTCGGAAAAACCCA	1480
Db	61	AACCTCTCAACGACCTCTACACAGATTTTGACACACTGACTGATTCCTCGGAAAAACCCA	120
QY	1490	TTTGTTTAAAGGTGGAGACATGTTGGTGACAAGTATATGACAGTGAAGTGGTTTACCAGAG	1540
Db	121	TTTGTTTAAAGGTGGAGACATGTTGGTGACAAGTATATGACAGTGAAGTGGTTTACCAGAG	180

QY	1479	GGAAAAACCCATTGTTTATTAAGGTGGAGACTGTTGGTGACAAGTATATGACAGTGAGTG	1538
Db	601	GGAAAAACCCATTGTTTATCAAGGTGGAACAGTGTGTGACAAGTATATGACAGTGAGTG	660
QY	1539	GTTTACCCAGAGCCATTCACCATGCACGATCCATCTGCCACCTGGCCTTGGACATGA	1598
Db	661	GCTTGCCAGAACCTTGTATCCACCATGCACGCTGTCATTTGCCCACTGGCTTTAGACATGA	720
QY	1599	TGGAATATCTGCTGCCAGGTTCAGTAGATGG	1629
Db	721	TGGAATATCTGCTGCTCAAGTTCANGTAGATGG	751
RESULT 15			
BI668844			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
1 (bases 1 to 745)			
NIH-MGC http://mgc.nci.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-r@mail.nih.gov			
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.			
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki			
Toshiyuki and Piero Carninci (RIKEN)			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLAM11796 row: a column: 03			
High quality sequence stop: 745.			
FEATURES			
source			
1. .745			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:5314058"			
/clone_lib="NIH_MGC_96"			
/tissue_type="hypothalamus"			
/lab_host="DH10B"			
/note="Organ: brain; Vector: pBluescriptR (modified			
pBluescript KS+); Site_1: BamHI; Site_2: Sali-XhoI (gtcgag			
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',			
size-selected for average insert size 2.3 kb and			
normalized to 500. This is a primary library enriched			
for full-length clones and constructed using the			
Cap-trapper method (Carninci, in preparation). Library			
constructed by M. Brownstein (NIH/NHGRI, National			
Institutes of Health). Note: this is a NIH_MGC Library."			
BASE COUNT			
ORIGIN			
Query Match 25.0%; Score 611.6; DB 13; Length 745;			
Best Local Similarity 98.3%; Pred. No. 3.5e-143;			
Matches 639; Conservative 0; Mismatches 9; Indels 2; Gaps 2;			
QY	166	AAAAGAGGACAGTATAGATGAAGAGGACAGTTCCTGTGCAAAATATATATGACATC	225
Db	91	AAAAGAGGACAGTATAGATGAAGAGGACAGTTCCTGTGCAAAATATATATGACATC	150
QY	226	CAAAACTTATGATTTGGTTGCTGCTGCAAGCAAGCTCAATCTCAATGCTGGAGAAAT	285
Db	151	CAAAACTTATGATTTGGTTGCTGCTGCAAGCAAGCTCAATCTCAATGCTGGAGAAAT	210

```
QY 286 CCTCCAAATGTTGGGAAGATGTTTTCGTCCTTTTGGCAAGAACTGCTTATGATACAAAT 345
Db 211 CCTCCAAATGTTGGGAAGATGTTTTCGTCCTTTTGGCAAGAACTGCTTATGATACAAAT 270
QY 346 CTTGGGTGCTCGGCTCTAATGTCAGAGAAATTTCTACAGAACCTTGATGCTCTGCACGA 405
Db 271 CTTGGGTGCTCGGCTCTAATGTCAGAGAAATTTCTACAGAACCTTGATGCTCTGCACGA 330
QY 406 CCACCTTGCTACCATCTACCCAGGAATCGGTGCACCTTCCCTTTAGGTGCACTGATGCAGA 465
Db 331 CCACCTTGCTACCATCTACCCAGGAATCGGTGCACCTTCCCTTTAGGTGCACTGATGCAGA 390
QY 466 AAAGGGCAAGAGACTCATTTTGCACCTACTACTCAGAGAGAGAGAGACTTCAGGATATTGT 525
Db 391 AAAGGGCAAGAGACTCATTTTGCACCTACTACTCAGAGAGAGAGAGACTTCAGGATATTGT 450
QY 526 CATTGGAATCATCAAAACAGTGGCACAAACAAATCCATGGCACTGAAATAGACATGAAGGT 585
Db 451 CATTGGAATCATCAAAACAGTGGCACAAACAAATCCATGGCACTGAAATAGACATGAAGGT 510
QY 586 TATTACGCAAGAAATGAAGAAATGTGATCATCTCAATTTTAAATTGAAGAAAGAGATC 645
Db 511 TATTACGCAAGAAATGAAGAAATGTGATCATCTCAATTTTAAATTGAAGAAAGAGATC 570
QY 646 AAAGAGAGAGATTTTATGAAGATCTTGACAGATTTGAAGAAA-ATGGTACCCAGGAAT 704
Db 571 AAAGAGAGAGATTTTATGAAGATCTTGACAGATTTGAAGAAACATGGTACCCAGGAAT 630
QY 705 CACGCATCAGCCCATATACATTTCTGAAAGCTTTTCCCTTTTCATATAATATTGACCGGG 764
Db 631 CACGCATCAGCCCATATACATTTCTGAAAGCTTTTCCCTTTTCATATAATATTGACCGGG 690
QY 765 ACCTAGTGGTCACTCAGTGTGGCAATGCTATATACAGAGTTCTCCCCCAG 814
Db 691 A-CTAGTGGTCACTCAGTGTGGCAATGCTATATACAGAGTTCTCCCCCAG 739
```

Search completed: July 1, 2003, 17:26:06
Job time : 2242.1 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 13:06:07 ; Search time 81.9108 Seconds
(without alignments)
9146.676 Million cell updates/sec

Title: US-09-762-767A-3
Perfect score: 2443
Sequence: 1 cccccccgcgcgtccgc.....aaatcatttttaaggagt 2443

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.4	3.3	3603	US-08-908-643C-85	Sequence 85, Appl
2	81.4	3.3	3645	US-08-908-643C-83	Sequence 83, Appl
3	81.4	3.3	3745	US-08-908-643C-84	Sequence 84, Appl
4	81.4	3.3	3787	US-08-908-643C-82	Sequence 82, Appl
5	60.2	2.5	7218	US-08-232-463-14	Sequence 14, Appl
6	56.4	2.3	4533	US-08-726-214-5	Sequence 5, Appl1
7	54.8	2.2	3784	US-07-623-033-1	Sequence 1, Appl1
8	53.6	2.2	3357	US-08-726-214-7	Sequence 7, Appl1
9	52.4	2.1	5199	US-08-726-214-13	Sequence 13, Appl
10	52.2	2.1	3518	US-09-412-210-2	Sequence 2, Appl1
11	50.8	2.1	4131	US-08-726-214-11	Sequence 11, Appl
12	50.6	2.1	1652	US-08-726-214-17	Sequence 17, Appl
13	50.6	2.1	4008	US-08-307-896-5	Sequence 5, Appl1
14	50.6	2.1	4008	US-08-726-214-3	Sequence 3, Appl1
15	50.6	2.1	4008	PCT-US95-11808-5	Sequence 5, Appl1
16	49.2	2.0	3924	US-08-726-214-9	Sequence 9, Appl1
17	48.6	2.0	2092	US-08-307-896-6	Sequence 6, Appl1
18	48.6	2.0	2092	PCT-US95-11808-6	Sequence 6, Appl1
19	48.6	2.0	4601	US-08-726-214-15	Sequence 15, Appl
20	47.4	1.9	7218	US-08-232-463-14	Sequence 14, Appl
21	46.2	1.9	3549	US-09-008-097-5	Sequence 5, Appl1
22	46.2	1.9	4342	US-09-474-076-1	Sequence 1, Appl1
23	46	1.9	4523	US-09-473-716-1	Sequence 1, Appl1
24	45.8	1.9	4046	US-07-793-961A-1	Sequence 1, Appl1
25	45.8	1.9	4046	US-08-240-357-1	Sequence 1, Appl1
26	42.6	1.7	4403765	US-09-103-840A-2	Sequence 2, Appl1
27	42.6	1.7	4411529	US-09-103-840A-1	Sequence 1, Appl1

C	28	42.4	1.7	19124	2	US-08-487-826B-13	Sequence 13, Appl
	29	42.2	1.7	658	4	US-08-998-416-595	Sequence 595, App
	30	40.4	1.7	1440	4	US-09-134-001C-1922	Sequence 1922, Ap
	31	40	1.6	51952	3	US-08-947-823-1	Sequence 1, Appl1
	32	39.6	1.6	747	4	US-09-134-001C-2248	Sequence 2248, Ap
	33	39.4	1.6	872	4	US-08-998-416-487	Sequence 487, App
	34	39.4	1.6	3701	4	US-08-845-258-10	Sequence 10, Appl
	35	39.4	1.6	3701	4	US-08-990-571-10	Sequence 10, Appl
	36	39.4	1.6	3701	4	US-08-723-142A-10	Sequence 10, Appl
	37	39.4	1.6	3701	4	US-09-528-784A-10	Sequence 10, Appl
	38	39.4	1.6	72604	4	US-09-268-992-7	Sequence 7, Appl1
	39	39.4	1.6	72604	4	US-09-657-474-7	Sequence 7, Appl1
	40	38.8	1.6	1422	1	US-08-319-704-5	Sequence 5, Appl1
	41	38.6	1.6	4473	3	US-08-894-173-1	Sequence 1, Appl1
	42	38.6	1.6	4473	4	US-09-398-193-1	Sequence 1, Appl1
	43	38.6	1.6	5361	4	US-08-973-462-2	Sequence 2, Appl1
	44	38.6	1.6	6152	4	US-08-973-462-1	Sequence 1, Appl1
	45	38.4	1.6	4946	3	US-08-817-188-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-908-643C-85
; Sequence 85, Application US/08908643C
; Patent No. 6120995
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; Pearlman, Joshua M.
; Barber, Michael T.
; Schultz, Stephanie
; Parkinson, Scott J.
; TITLE OF INVENTION: COMPOSITIONS THAT SPECIFICALLY BIND TO
; COLORECTAL CANCER CELLS AND METHODS OF
; USING THE SAME
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Woodcock Washburn Kurtz Mackiewicz & No. 6120995ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/908,643C
; FILING DATE: 07-Aug-1997
; CLASSIFICATION: N/A
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark Deluca
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-2209
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3603
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 85:
US-08-908-643C-85

Query Match 3.3% Score 81.4; DB 3; Length 3603;


```
QY 1359 TTAGTGGCATGTGGCTTCAATGCTTCTGTAGCAAGCATGCTATCGGAGAGAGCCCA 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2597 TCAGTGACATGTGAGTCTTCACTACTATCTGCAATATACACACCC-----CCA 2644
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1419 TGAAGATCGTCAACCTCCTCAGCAGCTCTACACAGATTTGACACACTGACTGATCC 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2645 TGAAGTGTGGACATGCTTAATGACATCTATAAGAGATTTTGACCAACATTTGTGATCATC 2704
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1479 GGAAGAACCCATTTGTTTATAAGTGGAGACTGTTGCTGACAGTATATGACAGTCAAGTG 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2705 ATGA-----TGCTACAGGTGGAAACCATCGGTGATGCTGATCGTGGCTAGTG 2755
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1539 GTTTACAGAGCCA---TGCATTCACCATGCAGATCCATCTGCCACCTGGCTTGGGA-- 1593
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2756 GTTTGCTTAAGAAATGCAATGCGCATGCAATAGACATTCGCAAGATGGCTTGGAAA 2815
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1594 -----CATGATGGAATGCTGGCCAGGTTCAGTAGATGCTGATGCTGATGATTC 2875
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2816 TCCTCAGCTTCATGGGGACCTTTGAGCTGGAGCATCTTCTGCGCTCCCAATATGATGATTC 2875
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1647 CAATAGGGATACACACTGGAGAGGTAGTTACAGGTCTCATAGGACAGCGGATGCCCTCGAT 1706
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2876 GCATTGGAGTTACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2935
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1707 ACTGCTCTTTTGGGAATGCTCAACCTCACAAGCCGACGAAACCAACAGGAGAAAGG 1766
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2936 ATTGCTATTGAGATACGGTCAACACAGCCCTCTAGGATGGAATCCACTGCGCTCCCTT 2995
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1767 GAAATAATGCTGCTGAATATAC 1791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2996 TGAGAATTCAGTGTGCTCCAC 3020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 5
US-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
```

```
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-Fls
US-08-232-463-14

Query Match 2.5%; Score 60.2; DB 1; Length 7218;
Best Local Similarity 5.6%; Pred. No. 6.9e-07;
Matches 23; Conservative 224; Mismatches 162; Indels 0; Gaps 0;

QY 1620 AAGTAGATGCTGAATCTGTTAGATAAACAATAGGAGATACACACTGGAGAGGTAGTTACAG 1679
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1431 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1372
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1680 GTGTCATAGACAGCGGATGCTCGATACCTGCTTTTGGGAATACCTGCAACCTCACAA 1739
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1371 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1740 GCGGACAGAACCAACAGAGAGAAAGGAGAAATAAATGTGCTGTAATATACATACAGAT 1799
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1311 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1800 GTCTTATGCTCCAGAAAATTCAGATCCCAATTCACATTCGACATGGAGACAGAGCCAGTGT 1859
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1251 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1860 CCATGAAGGCGCAAAAGAACCAATGCAAGTTGGTTTCTATCCAGAAAAATACAGGAA 1919
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1191 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1920 CAGAGAAACAAGCAGGATGATGACTGATCTGATGATCTGATGATGATGATGATGATGATGAT 1979
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1131 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1072
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1980 ACTAGGTCCAGTTTCTCTAACACAGCTGCCAAGCCAGGAGGAGGAGTCT 2028
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1071 RRRRRATCGCAAGCTCCCTCGACCTGCGACCAAGCTCGGAATTAATCT 1023
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 6
US-08-726-214-5
; Sequence 5, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,214
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
```


RESULT 11

US-08-726-214-11
; Sequence 11, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4131 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-726-214-11

Query Match 2.1%; Score 50.8; DB 3; Length 4131;
Best Local Similarity 56.6%; Pred. No. 0.00026;
Matches 94; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1638 TTCAGATACATAGGGGATACACTGGAGAGGTAGTTACAGGTCTCATAGCACAGCGGA 1697
DB 3291 TCCAGATGAAGATCGGGTGAACATGGGTCCGGTTGTAGCAGGTGTCTATTTGGGCGCCGGA 3350

QY 1698 TGCCTCGATACGTCTTTTGGGAATACCTGCAACCTCACAGCGGGAACACAGAACACAG 1757
DB 3351 AGCCACAGTATGACATCTCGGGAACACAGGTGAATGTTCCAGCCGTATGGACACAGCAG 3410

QY 1758 GAGAAAGGAAAAATAATGTCTGTAATATACATACAGATGTCT 1803
DB 3411 GAGTTCTTGACCGAATACAGGTGACCGAGATCTCTACCAAGTTCT 3456

RESULT 12

US-08-726-214-17
; Sequence 17, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1652 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-726-214-17

Query Match 2.1%; Score 50.6; DB 3; Length 1652;
Best Local Similarity 59.3%; Pred. No. 0.00017;
Matches 86; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1648 AATAGGGATACACACTGGAGAGGTAGTTACAGGTGTCTATAGCACAGCGGATGCTCGATA 1707
DB 1323 AGTGGGTATCAACCATGGGCTGTATAGTGGCGTCAATAGGGCTCAAAAGCCACAGTA 1382

QY 1708 CTGCTCTTTTGGGAATACCTCAACCTCACAGCGGGAACACACAGGAGAAAGG 1767
DB 1383 TGACATCTGGGCAACACTGTCAACGTGCCAGAGATGGACAGCCGGGTCTCTGGA 1442

QY 1768 AAAATAAATGTGTCTGAATATACA 1792
DB 1443 CAAATACAGTGTGACTGAGGAGACA 1467

RESULT 13

US-08-307-896-5
; Sequence 5, Application US/08307896C
; Patent No. 6034071
; GENERAL INFORMATION:
; APPLICANT: Iyengar, Srinivas Ravi
; TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND ADENYLYL
; TITLE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS
; FILE REFERENCE: 29770
; CURRENT APPLICATION NUMBER: US/08/307,896C
; CURRENT FILING DATE: 1994-09-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 4008
; TYPE: DNA
; ORGANISM: Rattus norvegicus

US-08-307-896-5

Query Match 2.1%; Score 50.6; DB 3; Length 4008;
Best Local Similarity 59.3%; Pred. No. 0.00029;
Matches 86; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1638 TTCAGATACATAGGGGATACACTGGAGAGGTAGTTACAGGTCTCATAGCACAGCGGA 1697
DB 3291 TCCAGATGAAGATCGGGTGAACATGGGTCCGGTTGTAGCAGGTGTCTATTTGGGCGCCGGA 3350

QY 1698 TGCCTCGATACGTCTTTTGGGAATACCTGCAACCTCACAGCGGGAACACAGAACACAG 1757
DB 3351 AGCCACAGTATGACATCTCGGGAACACAGGTGAATGTTCCAGCCGTATGGACACAGCAG 3410

QY 1758 GAGAAAGGAAAAATAATGTCTGTAATATACATACAGATGTCT 1803
DB 3411 GAGTTCTTGACCGAATACAGGTGACCGAGATCTCTACCAAGTTCT 3456

RESULT 14

US-08-726-214-17
; Sequence 17, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1652 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-726-214-17

Query Match 2.1%; Score 50.6; DB 3; Length 1652;
Best Local Similarity 59.3%; Pred. No. 0.00017;
Matches 86; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1648 AATAGGGATACACACTGGAGAGGTAGTTACAGGTGTCTATAGCACAGCGGATGCTCGATA 1707
DB 1323 AGTGGGTATCAACCATGGGCTGTATAGTGGCGTCAATAGGGCTCAAAAGCCACAGTA 1382

QY 1708 CTGCTCTTTTGGGAATACCTCAACCTCACAGCGGGAACACACAGGAGAAAGG 1767
DB 1383 TGACATCTGGGCAACACTGTCAACGTGCCAGAGATGGACAGCCGGGTCTCTGGA 1442

QY 1768 AAAATAAATGTGTCTGAATATACA 1792
DB 1443 CAAATACAGTGTGACTGAGGAGACA 1467

RESULT 15

US-08-307-896-5
; Sequence 5, Application US/08307896C
; Patent No. 6034071
; GENERAL INFORMATION:
; APPLICANT: Iyengar, Srinivas Ravi
; TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND ADENYLYL
; TITLE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS
; FILE REFERENCE: 29770
; CURRENT APPLICATION NUMBER: US/08/307,896C
; CURRENT FILING DATE: 1994-09-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 4008
; TYPE: DNA
; ORGANISM: Rattus norvegicus

US-08-307-896-5

Query Match 2.1%; Score 50.6; DB 3; Length 4008;
Best Local Similarity 59.3%; Pred. No. 0.00029;
Matches 86; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1638 TTCAGATACATAGGGGATACACTGGAGAGGTAGTTACAGGTCTCATAGCACAGCGGA 1697
DB 3291 TCCAGATGAAGATCGGGTGAACATGGGTCCGGTTGTAGCAGGTGTCTATTTGGGCGCCGGA 3350

QY 1698 TGCCTCGATACGTCTTTTGGGAATACCTGCAACCTCACAGCGGGAACACAGAACACAG 1757
DB 3351 AGCCACAGTATGACATCTCGGGAACACAGGTGAATGTTCCAGCCGTATGGACACAGCAG 3410

QY 1758 GAGAAAGGAAAAATAATGTCTGTAATATACATACAGATGTCT 1803
DB 3411 GAGTTCTTGACCGAATACAGGTGACCGAGATCTCTACCAAGTTCT 3456

Result No.	score	Query		Length	DB	ID	Description
		Match	%				
1	570	21.3	1302	9	US-10-106-698-961		Sequence 961, App
2	521	21.3	177556	9	US-09-952-213D-6		Sequence 6, App
3	197.4	8.1	15093	9	US-09-952-213D-5		Sequence 5, App
4	192.6	7.9	2850	9	US-10-205-823-163		Sequence 163, App
5	192.6	7.9	2871	9	US-10-205-823-164		Sequence 164, App
6	192.6	7.9	3017	9	US-10-205-823-161		Sequence 161, App
7	187.2	7.7	2430	9	US-09-952-213D-1		Sequence 1, App
8	147.2	6.0	2335	10	US-09-917-800A-1569		Sequence 1569, App
9	131.8	5.4	7697	9	US-09-952-213D-4		Sequence 4, App
10	130.6	5.3	177556	9	US-09-952-213D-6		Sequence 6, App
11	111.8	4.6	15093	9	US-09-952-213D-5		Sequence 5, App
12	86	3.5	955	9	US-09-952-213D-3		Sequence 3, App
13	81.4	3.3	3745	9	US-10-157-031-17		Sequence 17, App
14	81.4	3.3	3787	10	US-09-819-249-1		Sequence 1, App
15	56.4	2.3	7697	9	US-09-952-213D-4		Sequence 4, App
16	52.6	2.2	1037	9	US-10-198-846-13184		Sequence 13184, A
17	52.6	2.2	6196	9	US-10-282-942-1		Sequence 1, App
18	52.6	2.2	7053	10	US-09-764-864-792		Sequence 792, App
19	52.4	2.1	634	9	US-10-198-846-9332		Sequence 9332, App

QY 2054 GTCCTGTCCATTACCCCAAGACTTTCT

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.


```
Publication No. US20030108963A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Wonsay, Angela M.
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIORITY APPLICATION NUMBER: 60/307,982
PRIORITY FILING DATE: 2001-07-25
PRIORITY APPLICATION NUMBER: 60/314,356
PRIORITY FILING DATE: 2001-08-22
PRIORITY APPLICATION NUMBER: 60/325,020
PRIORITY FILING DATE: 2001-09-25
PRIORITY APPLICATION NUMBER: 60/341,746
PRIORITY FILING DATE: 2001-12-12
PRIORITY APPLICATION NUMBER: 60/362,158
PRIORITY FILING DATE: 2002-03-05
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 163
LENGTH: 2850
TYPE: DNA
ORGANISM: Homo sapiens
US-10-205-823-163

Query Match
Best Local Similarity 7.9%; Score 192.6; DB 9; Length 2850;
Matches 447; Conservative 0; Mismatches 334; Indels 24; Gaps 3;

QY 1004 CTCAGGCTCAAAATGATCTACTTACCTAGCAGATGATCTACTTTCTATGTTTCA 1063
DB 1481 CTCAGGCTCAAAATGATCTACTTACCTAGCAGATGATCTACTTTCTATGTTTCA 1540
QY 1064 AGTGTCATGAACCTGGAGCAATTTGACAGGAGGCGCTGATCTAAGTGACATCCCTCTG 1123
DB 1541 TGTGTGGACAGATTAGAAGATTTTACAGGACGAGGCTCTACCTCTCAGACATCCCAAT 1600
QY 1124 CATGATGCCAGCGGATCTGTTCTTTGGAGAACAAATTTAGAGAGGAATACAACTC 1183
DB 1601 CACAATGCATCGAGGATGTGGTCTTAATAGGGGAACAAGCCCGAGCTCAAGATGGCCTG 1660
QY 1184 ACCCAAGAACTGGAATCCTCACTGACAGGCTACAGCTCAGTTAAGAGCCCTGGAAGAT 1243
DB 1661 AAGAGAGCTGGGGAAGCTGAGCTACCTTGGAGAGCCCAAGCCCTGGAGAG 1720
QY 1244 GAAAGAAAGACACACATGCTGTATTTCTGCTTCCTCCCTGCTGTTGCCAATGAG 1303
DB 1721 GAGAGAAAGACATGAGACCTCTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAG 1780
QY 1304 CTGGGGCAGACGCTCAGTGGCTGCCAAGAGATGATGACATGTGACCATCCTCTTTAGT 1363
DB 1781 CTGTGGCAAGGCAAGTTGTGCAAGCCAAAGAGTTTCAGTAATGTCAACCATGCTCTCTCA 1840
QY 1364 GCAATGTGGGCTTCAATGCTTTCTGTAGCAGCATGATCTGGAGAGGAGCCATGAAG 1423
DB 1841 GACATCGTGGTTCACATGCCATCTGCTCCCAAGTCTC-----ACCGCTGCAG 1888
QY 1424 ATCGTCAACCTCCTCAACGACCTCTACACGAGTTTGACACACTGACTGATTTCCCGGAAA 1483
DB 1889 GTCATCACCATGCTCAATGCACTGTACACTGCTCTCGACACGACACTGTGGAGAGCTGGA- 1947
```

```
QY 1484 AACCCATTTGTTTATAGGTGGAGACTGTTGGTGACAGTATATACAGTGTGAGTTTGA 1543
DB 1948 -----TGCTACAAGGTGGAGACCATTTGGCGATGCTGCTATGTTGTAGCTGGGGATTA 1999
QY 1544 CCAGAGCCATGATTCACCATGCGACGATCCATCTGCCACCTGGCCCTTGGACATGATGGA 1603
DB 2000 CACAAGAGAGTGATGATCTATGCTGCTTCAGATAGCGCTGATGGCCCTGAAGATGATGGAG 2059
QY 1604 ATTGCTGCCAGGTT---CAAGTAGATGTTGAATCTGTTTCAGATAACAATAGGAGATACAC 1660
DB 2060 CTCCTGATGAAGTTATGCTCTCCCATGGAGAACCTATCAAGATCGGAATTTGAGCTGCAC 2119
QY 1661 ACTGGAGAGGTAGTTACAGGTGTCATAGACAGCGGATGCTCGATGACTGCTTTTGGG 1720
DB 2120 TCTGGATCAGTTTTTGGTGGCGTCTGTTGAGTTAAATGCCCGTTACTGCTGTTTTTGA 2179
QY 1721 AATAGTCAACCTCACAAGCGGCAACAGAACACAGAGAGAAAGGGAAGAAATAATATG 1780
DB 2180 AACATGTGCTCTGCTGCTAACAAATTTGAGTCTCCTGCTGCTACACAGAAATAATCATGTC 2239
QY 1781 TCTGAATATACATACAGATGCTTTA 1805
DB 2240 AGCCCAACAACCTTACAGATTACTCA 2264

RESULT 5
US-10-205-823-164
; Sequence 164, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsay, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIORITY APPLICATION NUMBER: 60/307,982
; PRIORITY FILING DATE: 2001-07-25
; PRIORITY APPLICATION NUMBER: 60/314,356
; PRIORITY FILING DATE: 2001-08-22
; PRIORITY APPLICATION NUMBER: 60/325,020
; PRIORITY FILING DATE: 2001-09-25
; PRIORITY APPLICATION NUMBER: 60/341,746
; PRIORITY FILING DATE: 2001-12-12
; PRIORITY APPLICATION NUMBER: 60/362,158
; PRIORITY FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164
; LENGTH: 2871
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-164

Query Match
Best Local Similarity 7.9%; Score 192.6; DB 9; Length 2871;
Matches 447; Conservative 0; Mismatches 334; Indels 24; Gaps 3;

QY 1004 CTCAGGCTCAAAATGATCTACTTACCTAGCAGATGATCTACTTTCTATGTTTCA 1063
DB 1502 CTCAGGCTCAAAATGATCTACTTACCTAGCAGATGATCTACTTTCTATGTTTCA 1063
```

QY 1064 AGTGTCATGAACCTGGACGATTTGACAAAGGAGGAGGCTGTATCTAAGTGACATCCCTCTG 1123
Db 1562 TGTGTGGACAGATTTAGAAATTTTACAGGACGAGGCTCTACCTCTCAGACATCCCAATT 1621
QY 1124 CATGATGCCACGGCATCTTGTCTTTTGGGAGAACATTTAGAGAGGAATACAACTC 1183
Db 1622 CACAATGCACTGAGGATGTGTCTTAATAGGGACAAAGCCGAGCTCAAGATGSCCTG 1681
QY 1184 ACCCAAGAACTGGAATCTCACTGACAGGCTACAGCTCACGTTAAGAGCCCTGGAAGAT 1243
Db 1682 AGAAGAGGCTGGGAGCTGAAGCTACCTTTGAGCAAGCCCAACCAAGCCCTGGAGGAG 1741
QY 1244 GAAAGAAAAGACAGACATTTGCTGTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1303
Db 1742 GAGAAGAAAAGACAGTACGCTTCTGTCTCCATATTTCCCTGTGAGGTGCTCAGCAG 1801
QY 1304 CTGGGCGCAAGGCTCCAGTGGCTGCCAAAAGATATGACAAATGTGACCATCTCTTTAGT 1363
Db 1802 CTGTGGCAAGGGCAAGTTGTGCAAGCCAAAGAAAGTTTCAAGTATGTACCATGTCTTCTCA 1861
QY 1364 GGCATTGTGGCTTCAATCTTCTGTAGCAAGCATGCTATCTGGAGAGGAGCCATGAAG 1423
Db 1862 GACATCGTTGGTTACTGCCATCTGCTCCAGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1909
QY 1424 ATGCTCAACCTCTCAACGACCTCTACACAGATTTGACACACTGACTGATTTCCCGGAAA 1483
Db 1910 GTCATCACCATGCTCAATGCACTGTACACTCGCTCGACGAGGAGTGTGGAGAGCTGGA- 1968
QY 1484 AACCCATTTGTTAAGGTGAGACTGTGTGACAAAGTATGACAGTATGACAGTGTGTTTA 1543
Db 1969 -----TGTCTAAGGTGGAGACCATTTGGCGATGCTTATTTGTAGCTGGGGGATTA 2020
QY 1544 CCAGAGCCATGCTACCATCAGCATCCATCTGCGACCTGGCCCTTGACATGATGGA 1603
Db 2021 CACAAGAGAGTGATCTATCTGTTTACAGTATGCTGATGCTGCTGATGCTGATGCTGAT 2080
QY 1604 ATGCTGGCAGGTTT-----CAAGTAGATGTGATCTGTTTCAAGTATGATGATGATGAT 1660
Db 2081 CTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2140
QY 1661 ACTGGAGAGTGTGTACAGTGTGTATAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1720
Db 2141 TCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2200
QY 1721 AATAGTGTCAACCTCACAAGCCGAAACCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1780
Db 2201 AACAATGTCACTGTGCTTAAACAAATTTGAGTCTGCTGAGTGTACCAAGGAGGAGGAGG 2260
QY 1781 TCTGTATATACATACAGTGTCTTA 1805
Db 2261 AGCCCAACAACTTACAGATTAATCTCA 2285

RESULT 6
US-10-205-823-161
; Sequence 161, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlengel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wansley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xunel
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044

; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 3017
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-205-823-161

Query Match 7.9%; Score 192.6; DB 9; Length 3017;
Best Local Similarity 55.5%; Pred. No. 2.3e-40;
Matches 447; Conservative 0; Mismatches 334; Indels 24; Gaps 3;

QY 1004 CTCAGGGTCAAATGATCTACTTACCTGAAGCAGATAGCATACTTTTCTATGTTCCACCA 1063
Db 1615 CTCAAAGGCAAATGATCTACTTGTGAATCCAGTGCATCTGTGTTTGGGTCACCC 1674
QY 1064 AGTGTCTGAACCTGGACGATTTGACAGGAGGAGGCTGTATCTAAGTACATCCCTCTG 1123
Db 1675 TGTGTGGACAGATTTAGAAATTTTACAGGAGGAGGCTCTACCTCTCAGACATCCCAAT 1734
QY 1124 CATGATGCCACGCGCATCTTCTTTTGGGAGAACATTTAGAGAGGAGGAGGAGGAGGAG 1183
Db 1735 CACATGCACTGAGGATGTGTCTTAATAGGGAACAGCCGAGCTCAAGATGCCCTG 1794
QY 1184 ACCCAAGAACTGGAATCTCTACTGACAGGCTACAGCTCACGTTAAGAGCCCTGGAGAT 1243
Db 1795 AAGAAGAGGCTGGGAGAGCTGAAGGCTACCTTTGAGCAAGCCCAACCAAGCCCTGGAGG 1854
QY 1244 GAAAGAAAAGACAGACATTTGCTGTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1303
Db 1855 GAGAAGAAAAGACAGATGAGCTTCTGTGCTCCATTTTCTCTGTGAGGTGCTCAGCAG 1914
QY 1304 CTGCGCACAAAGCTCCAGTGTGCTGCCAAAAGATATGACAATGTGACCATCTCTTTAGT 1363
Db 1915 CTGTGCGAAGGCAAGTTGTGCAAGCAAGAGATTCAGTAATGTACCATGTCTTCTCA 1974
QY 1364 GGCATTGTGGCTTCAATGCTTTCTGTAGCAAGCATGCTATCTGGAGAGGAGGAGGAG 1423
Db 1975 GACATCGTTGGTTCAGTCCCATCTGCTCCAGTGTCTCTCTCTCTCTCTCTCTCTCTCT 2022
QY 1424 ATCGTCAACCTCTCAAGGAGCTCTACACCATTTGACACACTGACTGACTGACTGACTG 1483
Db 2023 GTCATCACCATGCTCAATGCACTGTACACTCTCTCGACCAAGCATGTGTGAGAGGTGGA- 2081
QY 1484 AACCCATTGTTTATAAGGTGGAGACTGTGTGTGTGACAAAGTATATGACAGTGTGTTTA 1543
Db 2082 -----TGTCTACAAGGTGGAGACCATTTGGCGATGCTATTTGTGTAGCTGGGGATTA 2133
QY 1544 CCAGAGCCATGATTCACCATGACAGATCCATCTGCCACTGGCTTGGCATATGATGAGAA 1603
Db 2134 CACAAGAGAGTGATACACTCTGTTTCCAGATGAGGCTGATGGCCCTGAAGATGAGGAG 2193
QY 1604 ATTGCTGGCCAGGTTT-----CAAGTATAGTGTGATCTGTTTCAATACATACATAGGATAC 1660
Db 2194 CTCTGTGATGAGTATGATGTTCTCCCATGAGGAAACCTATCAAGATGCAATGGAGTGCAC 2253
QY 1661 ACTGGAGAGGTGTATACAGGTGTCTATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1720
Db 2254 TCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2313
QY 1721 AATAGTGTCAACCTCACAAGGCGGAGAACACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1780

Db 2314 AACATGTCACCTCTGGCTAACAAATTTGAGTCTCGAGTGTACACGAAATCAATGTC 2373
QY 1781 TCTGAATATACATACAGATGCTTGA 1805
Db 2374 AGCCCAACAACCTTACAGATTACTCA 2398

RESULT 7
US-09-952-213D-1
; Sequence 1, Application US/09952213D
; Publication No. US20030096240A1
; GENERAL INFORMATION:
; APPLICANT: MURAD, FERID
; APPLICANT: SHARINA, IRAIDA G.
; APPLICANT: KRUMENACKER, J. S.
; APPLICANT: MARTIN, E.
; TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC
; FILE REFERENCE: UTSH:252US
; CURRENT APPLICATION NUMBER: US/09/952,213D
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2430
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (286)..(2361)
US-09-952-213D-1

Query Match 7.7%; Score 187.2; DB 9; Length 2430;
Best Local Similarity 55.1%; Pred. No. 5.3e-39;
Matches 444; Conservative 0; Mismatches 338; Indels 24; Gaps 3;

QY 1003 TCTCAAGGCTCAATGATCTACTTACCTGAACACATAGCATACTTTTCTATGTCACC 1062
Db 1380 TCTCAAGGCTCAATGATCTACTTACCTGAACACATAGCATACTTTTCTATGTCACC 1439
QY 1063 AAGTGTCTGAACCTGGAGGATTTGACAGGAGAGGGCTGTATCTAAGTGACATCCCTCT 1122
Db 1440 ATGTGTGGACAGCTGGAAGATTTACAGGAGGGGGCTCTATCTGCGACATCCCAT 1499
QY 1123 GCATGATGCCAGCGGGATCTTGTCTTTTGGGAGAACAAATTTAGAGAGGAATACAACT 1182
Db 1500 TCATAACGCCCTGAGGGATGTTGTCTTGATAGGGGAGCAGGCGGGCTCAAGATGGCCT 1559
QY 1183 CACCAAGAATCTGGAATCTCTCACTGACAGGCTACAGCTCAGCTTAAGAGCCCTGGA 1242
Db 1560 CAAGAAGAGTTGGGAGCTGAAGCAACCTTGGAGCATGCCCAAGCCCTGGAGGA 1619
QY 1243 TGAAGAAGAAAGACAGACACATTTGCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCT 1302
Db 1620 AGAAGAAGAGAGGACAGTGGATCTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1679
QY 1303 GCTGGGCAACAGCGTCCAGTGGCTGCGCCAAAGATATGACATGTGACCATCTCTTTAG 1362
Db 1680 GCTGTGGCAAGACAAATTTGCAAGCCCAAGAAATTCAGGAGGTCAACCATGCTTTCTC 1739
QY 1363 TGGCATTTGGGCTTCAATGCTTTCTGTAGCAAGCATGATCTGGAGAAGGAGCCATGAA 1422
Db 1740 AGATATCGTAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1787
QY 1423 GATCGTCAACCTCTCAAGGCTCTTACACAGATTTGACACACTGACTGATTCGCGGAA 1482
Db 1788 GGTATCATCATGCTCAAGGCTCTTACACTGCTTTGACCAAGCTGTTGAGAGAGCTGGA 1847
QY 1483 AAACCCATTTGTTAAGGTGGAGACTGTTGGTGACAAAGTATATGACAGTGGAGTTT 1542
Db 1848 -----TGCTACAAAGGTGAGACCATCGGGATGATATTTGTGTGCGAGGTGAT 1898
QY 1543 ACCAGAGCCATGCATTCACCATGACAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1602

Db 1899 GCACAGAGAGAGTGACACCCATGCTGTCCAGATAGCACTGATGGCCCTGAAGATGATGA 1958
QY 1603 AATTGCTGCCAGG---TTCAAGTAGATGCTGAATCTGTTTCAGATAACAAATAGGGATACA 1659
Db 1959 GCTCTCCATGAGGTCATGCTCCCCACGAGAACCTATCAAGATCGGAATGGACTACA 2018
QY 1660 CACTGGAGAGGTAGTTTACAGGTGTATAGGACAGCGGATGCCCTGATGCTGTTTGG 1719
Db 2019 TTCTGGATCAGTGTGCTGGAGTGTCTCGAGTGAAGATGCCCGGTATTGCTGTTGG 2078
QY 1720 GAATAGTGTCAACCTCACAAGCCGACACAAACCCAGGAGAGAAAGGAAATAATGT 1779
Db 2079 AAACAATGTCACTGCTGCTTAACAAATTTGAATCTCGCTGCTGCTGCTGCTGCTGCT 2138
QY 1780 GTCTGAATATACATACAGATGCTTGA 1805
Db 2139 CAGCCCCACACATACAGGTTACTCA 2164

RESULT 8
US-09-917-800A-1569
; Sequence 1569, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1569
; LENGTH: 2335
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012770
US-09-917-800A-1569

Query Match 6.0%; Score 147.2; DB 10; Length 2335;
Best Local Similarity 51.0%; Pred. No. 2.4e-28;
Matches 517; Conservative 0; Mismatches 443; Indels 54; Gaps 5;

QY 725 TTCTCAAGAGCTTTTCTCTTTCATATAATTTAGCCGGACCTAGTGGTCACTCAGTGT 784
Db 771 TTCTGTGATGCTTTTCTCTTTCACATTTGCTTCCAGTGAAGCACTAAGGGTCAAGCAAGCT 830
QY 785 GGCATGCTATATACAGATGTTCTCCCGCAGCTCCAGCTTGGGAATTTGCAGCCCTTCGTCT 844

Db 831 GGAGTGAATTCAGAAATGTCCTCCGGAATCTTAACCCAGAAAGTTTGCACATAGATGAG 890
QY 845 GTCTTCGCTGGTTCCTCATATGATATAGTTTCCATGGATCTCTTCTCACATC 904
Db 891 TATTTTCCATCATCCACCTCAAGTTACTTTCAACATCTCCAGCATCTCAAGTTTCAT 950
QY 905 AATACGTGTTTGTATGAGAACCAAGGAAGGATGTTGGATGTGGAGAAATTAGAATGT 964
Db 951 AACAGTCAGTTTGTCTTGAGACAAGAAAGAAATGATGCC-----AAAGC 997
QY 965 GAGGATGAATGACTGGGACTGAGATCAGCTGCTTACGCTCAAGGTCAGAAATGATCTAC 1024
Db 998 AAGAAGAGCCACCCGATGCTCAA-----CTCCGGGTGATGATGATCTGG 1043
QY 1025 TTACCTGAAGCAGATAGCATCTTTTCTATGTTTCCAAAGTGTCTAAGGTCAGAAATGATCTAC 1084
Db 1044 ATGGAGTCTCTGAGGTGATGATCTTCATGTTTCCCAACGCTCCGACGCTGCAAGAG 1103
QY 1085 TTGACAAGAGAGGGCTGTATCTAAGTGACATCCCTCTGATGATGATGATGATGATGAT 1144
Db 1104 CTGGAAGAGAGCAAGATGATCTTTCTGATATCGCTCCGACGACGACGACGAGGATCTC 1163
QY 1145 GTTCTTTTGGGAACAATTTAGAGAGGAATAACAATCTACCCAAGAACTGGAAATCCCTC 1204
Db 1164 ATCTCTCAACAGCAGAGGCTGGCAGAGATGAGCTGTCTGCTCCCACTTGGAAAGAAG 1223
QY 1205 ACTGACAGGCTACAGCTCAGTTAAGAGCCCTGGAAGATGAAAGAAAGACAGACACA 1264
Db 1224 AAGGAGAGTTGGTGTCTTCCATCACTGGCCATCGAAGAAGAAGACAGAGAC 1283
QY 1265 TTGCTGTATCT 1324
Db 1284 TTGCTGTATGCTATGCTGCTGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1343
QY 1325 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1384
Db 1344 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1403
QY 1385 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1444
Db 1404 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1451
QY 1445 CTCTACACAGATTTACACACTGACTGATCTCCGGAAGAAACCCATTTGTTTATAGGTT 1504
Db 1452 ATGACTCCAGTTTGACAGGTTTAAACAGTGTCTCATGA-----TGCTCAAGTA 1502
QY 1505 GAGACTGTTGGTGACAAGTATAGCAGTGTGAGTGTGTACAGAGCCATGCAATTCACCAT 1564
Db 1503 GAAACAATAGGGATGCTTACATGCTGTTGGTGGAGTACCACTACCCGTTGAAAGCCAT 1562
QY 1565 GCAGATCCATCTGCCACCTGGCCTTGGACATGA-----TGGAAATGCTGCCAGGTT 1618
Db 1563 GCTCAAGAGTCCCAATTTTGTCTGGGATGAGAATTTCTCAAAAGAGTGAAT 1622
QY 1619 CAAGTAGATGCTGAATCTGTTCAAGATAACATAGGATAGATACACTGGAGAGGTAGTACA 1678
Db 1623 CCGTCACTGGGAAGCACTATCCAGATCAGAGTGGGAATCCACACTGGACCACTCTTAGCA 1682
QY 1679 GGTGTATAGACAGAGGATGCTCTGATCTGCTTTTGGGAATCTGCTCAAC 1732
Db 1683 GGTGTGTGGGACAGAGATGCTCTGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1736

RESULT 9

US-09-952-213D-4

; Sequence 4, Application us/09952213D

; Publication No. US20030096240A1

; GENERAL INFORMATION:

; APPLICANT: MURAD, FERID

; APPLICANT: SHARINA, IRAIDA G.

; APPLICANT: KRUMENACKER, J. S.

; APPLICANT: MARTIN, E.

; TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC

; FILE REFERENCE: UTSH:252US

; CURRENT APPLICATION NUMBER: US/09/952,213D

; CURRENT FILING DATE: 2002-08-16

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 7697

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: modified_base

; LOCATION: (605)..(6955)

; OTHER INFORMATION: N = A, C, T/U OR G

; US-09-952-213D-4

Query Match 5.4%; Score 131.8; DB 9; Length 7697;

Best Local Similarity 56.4%; Pred.No. 7.3e-24;

Matches 286; Conservative 3; Mismatches 205; Indels 13; Gaps 2;

QY 1003 TCTCAAGGGTCAAAATGATCTACTTACCTGAAGCAGATAGCATATCTTTTCTAAGTGCACCTCT 1062

Db 4262 TCTCAAGGGTCAAAATGATCTACTTACCTGAAGCAGATAGCATATCTTTTCTAAGTGCACCTCT 4321

QY 1063 AAGTGTCAATGAACCTGACGATTTGACAGGAGAGGCTGTATCTAAGTGCACCTCTCT 1122

Db 4322 ATGTGTGGACAGGCTGGAAGATTTACAGGACGGGGCTCTATCTGTCCGACATCCCAAT 4381

QY 1123 GCATGATGCCACCGCGATCTTGT-TCTTTTGGAGAACAAATTTAGAGAGGAATACAAAC 1181

Db 4382 TCATAACGCCCTGAGGGATGTTTTCATGATAGGGGAGAGGACCGGGCTCAAGATGGCC 4441

QY 1182 TCACCAAGAACCTGGAATCTCTCACTGACAGGCTACAGTCTACGTTAAGAGCCCTCGAAG 1241

Db 4442 TCAAGAAGAGGTTGGGGAAGCTGAAGCAACCTGGAGCATGCCCAAGCCCTGGAGG 4501

QY 1242 ATGAAAGAAAGACAGACACATCTGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1301

Db 4502 ARGAGAAGAGGACAGCATGATGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4561

QY 1302 AGCTGGGGCACAAGCTCCAGTCCCTGCAAAAGATATGACAAATGTGACCATCTCTTTA 1361

Db 4562 AGCTGTGGCAGACAAATTTGCAAGCCCAAGAAATTCAGGAGGTACCATCTCTTTCT 4621

QY 1362 GTGGCATGTGGGCTTCAATGCTTTCTGTAGCAAGCATGCTGTGGAGAGGAGCCATGA 1421

Db 4622 CAGATATGTAGGGTTCACTGCTATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4669

QY 1422 AGATCTCAACCTCCCTCAAGGACCTCTACACAGATTTGACACACTGACTGATCCCGGA 1481

Db 4670 AGGTATCATCGATGCTCAACGCTCTCTACACTCGCTTTTGACGAGGTGGAGCTGG 4729

QY 1482 AAAAAACCAATTTGTTTATAAGGTGGAGA 1508

Db 4730 ATCTCTACAAGGTAGGGAAGGTGAAA 4756

RESULT 10

US-09-952-213D-6/c

; Sequence 6, Application us/09952213D

; Publication No. US20030096240A1

; GENERAL INFORMATION:

; APPLICANT: MURAD, FERID

; APPLICANT: SHARINA, IRAIDA G.

; APPLICANT: KRUMENACKER, J. S.

; APPLICANT: MARTIN, E.

; TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC

; FILE REFERENCE: UTSH:252US

; CURRENT APPLICATION NUMBER: US/09/952,213D

; CURRENT FILING DATE: 2002-08-16

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 177556


```
; LENGTH: 3745
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-157-031-17

Query Match      3.3%; Score 81.4; DB 9; Length 3745;
Best Local Similarity 52.4%; Pred. No. 1.2e-10;
Matches 296; Conservative 0; Mismatches 236; Indels 33; Gaps 4;

QY 1239 AAGATGAAAAGAAAAGACAGACACATTCCTGCTATTTCTGCTCTCCCTGCTGTTGCCA 1298
DB 2408 AGCAGAGAGGACAGAGGCTGACAGACTTAATTTATGTTGTTTCCAGGCTAGTGGTAA 2467
QY 1299 ATGAGCTGGGACACAGCTCCAGTCCGCTGCCAAAAGATATGACAAATGACCACTCTCT 1358
DB 2468 AGTCTCTGAAGGAGAAAGGCTTTGTGGAGCCGGAACATATAGGAAGTTTACAATCTACT 2527
QY 1359 TTATGGCAATTTGGGCTTCAATGCTTTCTGTAGCAGCATGCATCTGGAGAGGAGCCA 1418
DB 2528 TCAGTGACATTTAGGTTTCACTACTATCTGCAATACAGCACCC-----CCA 2575
QY 1419 TGAAGATCGTCAACCTCTCAACGACCTCTACACACAGATTTGACACACTGACTGATCC 1478
DB 2576 TGGAGTGGTGGACATGCTTAATGACATCTATAAGAGTTTTCACACATTTGATCATC 2635
QY 1479 GGAATAACCCATTTGTTTATAAGTGGAGACTGTTGGTGACAAAGATATATGACAGTAG 1538
DB 2636 ATGA-----TGCTACAAGGTGGAACCATCGGTGATCGGTACATGGTGGCTAGTG 2686
QY 1539 GTTTACCAGAGCCA---TGCAATCACCATGACGATCCATCTGCCACCTGGCCTTGA-- 1593
DB 2687 GTTTCCTAAGAGAAATGCGCATGCGCAATAGCAATGCCAAGATGGCCTTGGAAA 2746
QY 1594 -----CATGATGAAATTTGCTGCCAGGTTCAAGTAGATGTTGAATCTGTTCAAGATA 1646
DB 2747 TCCTCAGCTTCATGGGACCTTTGAGCTGGAGCACTTCTTGCCCTCCCAATATGATGATC 2806
QY 1647 CAATAGGATACACACTGAGAGGTAGTTTACAGGTGTCATAGGACAGCGGATGCCTCGAT 1706
DB 2807 GCATTTGGAGTTCACTCTGCTGCTCCCTGCTGCTGGAGTTTGGGAATCAAGATGCCTCGT 2866
QY 1707 ACTGCTCTTTTGGGAATCTGTCACCTCACAAGCCGGAACACACAAACCACAGGAGAAAG 1766
DB 2867 ATGCTATTTGGAGATACGGTCAACAGCCTCTAGGATGGATCCACTGGCCTCCCTT 2926
QY 1767 GAAAATAAATGTGCTGAATATAC 1791
DB 2927 TGAGAAATTCACGTGAGTGGCTCCAC 2951

RESULT 14
US-09-819-249-1
; Sequence 1, Application US/09819249
; Patent No. US20010029019A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; APPLICANT: Park, Jason
; APPLICANT: Schulz, Stephanie
; TITLE OF INVENTION: Compositions And Methods For Identifying And Targeting Cancer Cell
; FILE REFERENCE: TJU2412
; CURRENT APPLICATION NUMBER: US/09/819,249
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,229
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3787
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (118)..(3336)
US-09-819-249-1

Query Match      3.3%; Score 81.4; DB 10; Length 3787;
Best Local Similarity 52.4%; Pred. No. 1.2e-10;
Matches 296; Conservative 0; Mismatches 236; Indels 33; Gaps 4;

QY 1239 AAGATGAAAAGAAAAGACAGACACATTCCTGCTATTTCTGCTCTCCCTGCTGTTGCCA 1298
DB 2477 AGCAGAGAGGACAGAGGCTGACAGACTTAATTTATGTTGTTTCCAGGCTAGTGGTAA 2536
QY 1299 ATGAGCTGGGACACAGCTCCAGTCCGCTGCCAAAAGATATGACAAATGACCACTCTCT 1358
DB 2537 AGTCTCTGAAGGAGAAAGGCTTTGTGGAGCCGGAACATATAGGAAGTTTACAATCTACT 2596
QY 1359 TTATGGCAATTTGGGCTTCAATGCTTTCTGTAGCAGCATGCATCTGGAGAGGAGCCA 1418
DB 2597 TCAGTGACATTTAGGTTTCACTACTATCTGCAATACAGCACCC-----CCA 2644
QY 1419 TGAAGATCGTCAACCTCTCAACGACCTCTACACACAGATTTGACACACTGACTGATCC 1478
DB 2645 TGGAGTGGTGGACATGCTTAATGACATCTATAAGAGTTTTCACACATTTGATCATC 2704
QY 1479 GGAATAACCCATTTGTTTATAAGTGGAGACTGTTGGTGACAAAGATATATGACAGTAG 1538
DB 2705 ATGA-----TGCTACAAGGTGGAACCATCGGTGATCGGTACATGGTGGCTAGTG 2755
QY 1539 GTTTACCAGAGCCA---TGCAATCACCATGACGATCCATCTGCCACCTGGCCTTGA-- 1593
DB 2756 GTTTCCTAAGAGAAATGCGCATGCGCAATAGCAATGCCAAGATGGCCTTGGAAA 2815
QY 1594 -----CATGATGAAATTTGCTGCCAGGTTCAAGTAGATGTTGAATCTGTTCAAGATA 1646
DB 2816 TCCTCAGCTTCATGGGACCTTTGAGCTGGAGCACTTCTTGCCCTCCCAATATGATGATC 2875
QY 1647 CAATAGGATACACACTGAGAGGTAGTTTACAGGTGTCATAGGACAGCGGATGCCTCGAT 1706
DB 2876 GCATTTGGAGTTCACTCTGCTGCTCCCTGCTGCTGGAGTTTGGGAATCAAGATGCCTCGT 2935
QY 1707 ACTGCTCTTTTGGGAATCTGTCACCTCACAAGCCGGAACACACAAACCACAGGAGAAAG 1766
DB 2936 ATGCTATTTGGAGATACGGTCAACAGCCTCTAGGATGGATCCACTGGCCTCCCTT 2995
QY 1767 GAAAATAAATGTGCTGAATATAC 1791
DB 2996 TGAGAAATTCACGTGAGTGGCTCCAC 3020

RESULT 15
US-09-952-213D-4/c
; Sequence 4, Application US/09952213D
; Publication No. US20030096240A1
; GENERAL INFORMATION:
; APPLICANT: MURAD, FERID
; APPLICANT: SHARINA, IRAIDA G.
; APPLICANT: KRUMENACKER, J. S.
; APPLICANT: MARTIN, E.
; TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC
; FILE REFERENCE: UTSH:252US
; CURRENT APPLICATION NUMBER: US/09/952,213D
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 7697
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (605)..(6955)
; OTHER INFORMATION: N = A, C, T/U OR G
US-09-952-213D-4
```

```
Query Match      2.3%; Score 56.4; DB 9; Length 7697;
Best Local Similarity 53.1%; Pred.No. 0.00095;
Matches 120; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 1638 TTCAGATAACAATAGGGATACACACTGGAGAGGTAGTTACAGGTGTCATAGGACAGCGGA 1697
    |||||  |||  ||  ||||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 6512 TTCAGATCGAATTGGACTACATTCGGATCAGTGTCTGGAGTTCGCGAGTGAAGA 6453
    |||||  |||  ||  ||||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY 1698 TGCCTCGATACGTCTTTTGGGAATAGTCAACCTCACAGCCGAAACAGAACCCACAG 1757
    ||||  ||  ||  ||  |||||  ||  |||||  ||  ||  ||  ||  ||  ||
Db 6452 TGCCTCGATACGTCTTTTGGGAATAGTCAACCTCACAGCCGAAACAGAACCCACAG 6393
    ||||  ||  ||  ||  |||||  ||  |||||  ||  ||  ||  ||  ||  ||
QY 1758 GAGAAAGGGAAATAATGTCTGATATACATACAGATGCTTATGCTCCAGAAA 1817
    ||  ||  |||||  |||||  |||||  |||||  |||||  ||  ||  ||  ||  ||
Db 6392 GTGTGCTCGGAAATCAATGTCAGCCCCACACATACAGGTATGGATGGCACTAAGGAC 6333
    ||  ||  |||||  |||||  |||||  |||||  |||||  ||  ||  ||  ||  ||
QY 1818 ATTCAAGATCCCAATTCACCTTGGAGCACAGAGGCCAGTGTCAT 1863
    ||  ||  ||  |||||  ||  |||||  ||  |||||  ||  |||||  ||
Db 6332 CTGAATGCATCATCTCACTTGACAACTTGTCATCTTAGTCCAT 6287
    ||  ||  ||  |||||  ||  |||||  ||  |||||  ||  |||||  ||
```

Search completed: July 1, 2003, 20:37:28
Job time : 237.304 secs

JOURNAL antibodies, and for somatic gene therapy of arteriosclerosis
Patent: DE 19837015-A 1 24-FEB-2000;
VASOPHARM BIOTECH GMBH & CO KG (DE)
FEATURES Location/Qualifiers
source 1..3015
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 877 a 680 c 716 g 742 t
ORIGIN

Query Match		100.0%;	Score 3015;	DB 6;	Length 3015;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 3015;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CCCTATATGGGATTTGGCGGCTGCAGAGACACGAGCTCAGTTCCTCCCTGCGCTAGTCAG	60		
Db	1	CCCTATATGGGATTTGGCGGCTGCAGAGACACGAGCTCAGTTCCTCCCTGCGCTAGTCAG	60		
Qy	61	CTTAGTGGTGGGACTCAGCTCAGAGTCAGTTCAGAGAGAGAGTTCAGTGCAGAGTT	120		
Db	61	CTTAGTGGTGGGACTCAGCTCAGAGTCAGTTCAGAGAGAGAGTTCAGTGCAGAGTT	120		
Qy	121	TTCTACACATTTTCTCGCTAGAGCAGCAGAGCTGGAACAGACCCAGGCGGAGAC	180		
Db	121	TTCTACACATTTTCTCGCTAGAGCAGCAGAGCTGGAACAGACCCAGGCGGAGGAC	180		
Qy	181	ACCTGTGGGAGGAGCGCTGGAGAGCTTAGAGACCCCGCGGCTGATCTCACC	240		
Db	181	ACCTGTGGGAGGAGCGCTGGAGAGCTTAGAGACCCCGCGGCTGATCTCACC	240		
Qy	241	ATGTGGGATTTGGAGGGCGGCTGGAGCTGTAGAGTCCGGAAGACAGCCCGGAG	300		
Db	241	ATGTGGGATTTGGAGGGCGGCTGGAGCTGTAGAGTCCGGAAGACAGCCCGGAG	300		
Qy	301	GTGTGCAAGCACCAGACTGCGGCTTTGGAGAAAGCGTGAAGAGGCGGCGACCGCG	360		
Db	301	GTGTGCAAGCACCAGACTGCGGCTTTGGAGAAAGCGTGAAGAGGCGGCGACCGCG	360		
Qy	361	TCTCCGCGCTGTGTCACCGCTGCGCTGAGCTGCATATAGAACTACAG	420		
Db	361	TCTCCGCGCTGTGTCACCGCTGCGCTGAGCTGCATATAGAACTACAG	420		
Qy	421	TTACCACTGCTTGAATGATAGTGGCTTCTGTTGCTCAGTCTCATATAGAACTACAG	480		
Db	421	TTACCACTGCTTGAATGATAGTGGCTTCTGTTGCTCAGTCTCATATAGAACTACAG	480		
Qy	481	CTCATCAGGAGGATCGCAGAGGTAAAGACACCAACACCATGTTCTGCAGAGCT	540		
Db	481	CTCATCAGGAGGATCGCAGAGGTAAAGACACCAACACCATGTTCTGCAGAGCT	540		
Qy	541	CAAGGATCTCAAGATCAGAGAGTGTCTTTCTTACTGCGACCAAGTTC	600		
Db	541	CAAGGATCTCAAGATCAGAGAGTGTCTTTCTTACTGCGACCAAGTTC	600		
Qy	601	TAAGCACTCTCAGAGGAGCAGCAAGCTCAGAGAGCTGCAAAACCAACCGTGCCAT	660		
Db	601	TAAGCACTCTCAGAGGAGCAGCAAGCTCAGAGAGCTGCAAAACCAACCGTGCCAT	660		
Qy	661	CTGTCAAGACATCTCTGAGAGACATACAGAAAGTCTTCTCCTCAAGAAACCAAGTCG	720		
Db	661	CTGTCAAGACATCTCTGAGAGACATACAGAAAGTCTTCTCCTCAAGAAACCAAGTCG	720		
Qy	721	GAGCGGATCTATCTTACACTTTGGCAGAGAGTATTTGCAAACTGATTTCCAGAGTT	780		
Db	721	GAGCGGATCTATCTTACACTTTGGCAGAGAGTATTTGCAAACTGATTTCCAGAGTT	780		
Qy	781	TGAACGGCTGAATCTTCACTTACAGAAACATTTGGCAAGCACAACCAATAAAGAACAG	840		
Db	781	TGAACGGCTGAATCTTCACTTACAGAAACATTTGGCAAGCACAACCAATAAAGAACAG	840		
Qy	841	GAATCTTTTGAAGAGAGACTTTGAAACCAATTTGAGAGCAAGCTTGCAGCAGG	900		
Db	841	GAATCTTTTGAAGAGAGACTTTGAAACCAATTTGAGAGCAAGCTTGCAGCAGG	900		

Qy	901	AGTTCCAGTGGAGGTTATCAAGAATCTCTTGGTGAAGAGGTTTTTAAATATGTTACGA	960		
Db	901	AGTTCCAGTGGAGGTTATCAAGAATCTCTTGGTGAAGAGGTTTTTAAATATGTTACGA	960		
Qy	961	GGAAGATGAAAACATCCTTGGGGTGGTGGAGGACCCCTTAAAGATTTTTTAAACAGCTT	1020		
Db	961	GGAAGATGAAAACATCCTTGGGGTGGTGGAGGACCCCTTAAAGATTTTTTAAACAGCTT	1020		
Qy	1021	CAGTACCCCTTCTGAACACAGAGCAGCCATTGCCAGAGAGGAGGAGGAGGCGGCTTGA	1080		
Db	1021	CAGTACCCCTTCTGAACACAGAGCAGCCATTGCCAGAGAGGAGGAGGAGGCGGCTTGA	1080		
Qy	1081	GGAGCGCTCCATCTTATGCGCTGGATAAGAGGAGTATTTTCTACATGTTTACTTCTT	1140		
Db	1081	GGAGCGCTCCATCTTATGCGCTGGATAAGAGGAGTATTTTCTACATGTTTACTTCTT	1140		
Qy	1141	CCCTAAGAGAACACCTTCCCTGATTTCTCCGGCATCATAAAGGAGCTGCTCACGTATT	1200		
Db	1141	CCCTAAGAGAACACCTTCCCTGATTTCTCCGGCATCATAAAGGAGCTGCTCACGTATT	1200		
Qy	1201	ATATGAACGGAAGTGGAGTGTGTTAATGCCCTCCCTGCTTCCATATGATTGACGCGA	1260		
Db	1201	ATATGAACGGAAGTGGAGTGTGTTAATGCCCTCCCTGCTTCCATATGATTGACGCGA	1260		
Qy	1261	GTTTGTGAATCAGCCCTACTTGTGTACTCCCTTACATGAAAGACACCAAGCCATCCCT	1320		
Db	1261	GTTTGTGAATCAGCCCTACTTGTGTACTCCCTTACATGAAAGACACCAAGCCATCCCT	1320		
Qy	1321	GTCCCCAGAAACCCAGTCCCTGCTGGTGGATTCACATCGCTATTTGCAAGACATT	1380		
Db	1321	GTCCCCAGAAACCCAGTCCCTGCTGGTGGATTCACATCGCTATTTGCAAGACATT	1380		
Qy	1381	TCCATTCCATTTTCATGTTTGCACAAAGATATGACAAATTTGCAATTTGGCAATGCAAT	1440		
Db	1381	TCCATTCCATTTTCATGTTTGCACAAAGATATGACAAATTTGCAATTTGGCAATGCAAT	1440		
Qy	1441	AAGCTGTATGAACAGGAGAGACTTCAAGGAAAGCCCTTAATTTTGAAGAACTTTGAAAT	1500		
Db	1441	AAGCTGTATGAACAGGAGAGACTTCAAGGAAAGCCCTTAATTTTGAAGAACTTTGAAAT	1500		
Qy	1501	TCTGACTCCAAAATCAACCAGACCTTTACGGGATCATGCTATGTTGAATGCAAGTT	1560		
Db	1501	TCTGACTCCAAAATCAACCAGACCTTTACGGGATCATGCTATGTTGAATGCAAGTT	1560		
Qy	1561	TGTTGTACGAGTGAAGAGATGGGACAACTCTGTGAAGAAATCTTCAAGGCTTATGGACCT	1620		
Db	1561	TGTTGTACGAGTGAAGAGATGGGACAACTCTGTGAAGAAATCTTCAAGGCTTATGGACCT	1620		
Qy	1621	CAAGGCGCAATGATCTACATTTGTAATCCAGTGCATCTTGTGTTGGGGTCAACCTG	1680		
Db	1621	CAAGGCGCAATGATCTACATTTGTAATCCAGTGCATCTTGTGTTGGGGTCAACCTG	1680		
Qy	1681	TGTGGCAGATTTAGAGATTTTACAGGACGAGGCTTCTACCTCTCAGACATCCCAATCA	1740		
Db	1681	TGTGGCAGATTTAGAGATTTTACAGGACGAGGCTTCTACCTCTCAGACATCCCAATCA	1740		
Qy	1741	CAATGCACTCAGGGATGTGGTCTTAAATAGGGGAACAGCCGAGCTCAAGTGCCTGAA	1800		
Db	1741	CAATGCACTCAGGGATGTGGTCTTAAATAGGGGAACAGCCGAGCTCAAGTGCCTGAA	1800		
Qy	1801	GAAGAGCTGGGGAAGCTGAAGCTACCCCTTGAAGAGCCCAAGCCCTGGAGGAGGA	1860		
Db	1801	GAAGAGCTGGGGAAGCTGAAGCTACCCCTTGAAGAGCCCAAGCCCTGGAGGAGGA	1860		
Qy	1861	GAAGAAAGACAGTACAGCTTCTGCTCCATTTTCCCTGTGAGTTCTCAGCAGCT	1920		
Db	1861	GAAGAAAGACAGTACAGCTTCTGCTCCATTTTCCCTGTGAGTTCTCAGCAGCT	1920		
Qy	1921	GTGGCAAGGCAAGTTGTGCAAGCCCAAGAGTTTCAATGTGTCACCATGCTCTTCTCAGA	1980		
Db	1921	GTGGCAAGGCAAGTTGTGCAAGCCCAAGAGTTTCAATGTGTCACCATGCTCTTCTCAGA	1980		

[illegible]

CDS	523...2676	/gene="GC-S-alpha-3"	
		/EC number="4.6.1.2"	
		/note="soluble guanylate cyclase large subunit"	
		/codon_start=1	
		/product="guanylate cyclase"	
		/protein_id="CAA47145.1"	
		/db_xref="GI:31684"	
		/translation="MFCITKLKDKITGCEPCFSLAPQGVNSESRAAGSSSEKATV	
		/db_xref="SWISS-PROT:Q02108"	
		PIQDPIPEKNIQSLPQKTSRSRVYLHTLAESICKLIPPEFERLNVALORTLAKHI	
		KESKSLERDEFKTAQAVQSPVELSKNLVLFKLVYTRKMTSLGWLEAPLKI	
		FKQLQPSEREDEPLPSRKGLQDASILCDKDDFLHVIYFFPFRKTSLLPLGLIK	
		AAAHLYTEVPSLMPDPCPHNCSFNQPYLLYSVHKMSTKPSLSPKQSSLVIP	
		TSLSFCTFPFHEFMDMTILOFNGIRLRLMNRDFQGNFNFYFELLTKINQTSFG	
		ITMLNNQFVVRVWDNSVKKSRYMDLQKMIYIVESSAILFLGSPCDRLDEDFTG	
		RGLVLDIPILHNALRDVLIGEORADQGLKRLGLIKATLEOAHOLEBEKKKVDLI	
		LSISFCEVAQIQQGVVQAKFESNVNMLFSDIVGFTALCSQCSPLQVITMLNLYT	
		RFDQCCGELDYIVETAMPVIMGLGHKESDTHAVOIALMALKMELSDVEMSPHGE	
		PIKMRIGLHSGVYFAGVYKMPRYCLFGNNVTLANKFESCVPKRNINVSPTTYRLIK	
		DCPQFVTPRSRELPNFPSEIPGICHFLDAYQQGTNSRQPCPKKDVEDASQFFRQS	
		IRNRLATYIPYIKSLGFDLSLKMCRASESTLGLVDG"	
BASE COUNT	874 a	679 c	710 g 741 t
ORIGIN			
Query Match	94.78;	Score 2855.4;	DB 9; Length 3004;
Best Local Similarity	98.6%;	Pred. No. 0;	
Matches 2976;	Conservative 0;	Mismatches 26;	Indels 15; Gaps 9;
QY	1	CCCTTAGCGGATGGCGGCTCGAGAGACCAGGACTCAGTTCCTCCCTGCCCTAGTCTGAG	60
Db	1	CCCTTAGCGGATGGCGGCTCGAGAGACCAGGACTCAGTTCCTCCCTGCCCTAGTCTGAG	60
QY	61	CCTAGTGGGTGGGACTCAGCTCAGAGTCAGTTCCTCAGAGAGAGGTTTCAGTGCAGAGTT	120
Db	61	CCTAGTGGGTGGGACTCAGCTCAGAGTCAGTTCCTCAGAGAGAGGTTTCAGTGCAGAGTT	120
QY	121	TTCTACACTTTTCTCGGCTAGAGCAGCAGAGCCCTGGAACAGACCCAGGCGGAGGAC	180
Db	121	TTCTACACTTTTCTCGGCTAGAGCAGCAGAGCCCTGGAACAGACCCAGGCGGAGGAC	180
QY	181	ACCTGTGGGGAGGAGCGCTGGAGAGCTTAGAGACCCAGCCGCGGCGTGATCTCACC	240
Db	181	ACCTGTGGGGAGGAGCGCTGGAGAGCTTAGAGACCCAGCCGCGGCGTGATCTCACC	240
QY	241	ATGTGCGGATTTGGAGCGCGCTCGGAGCTGTAGAGATCCGGAAGACACAGCCCGGAG	300
Db	241	ATGTGCGGATTTGGAGCGCGCTCGGAGCTGTAGAGATCCGGAAGACACAGCCCGGAG	300
QY	301	GTGTGCGAAGCCAAAGACTGCGGCTTTGGAGAAAGCGTGAGAGGGGGCCACCGCGG	360
Db	301	GTGTGCGAAGCCAAAGACTGCGGCTTTGGAGAAAGCGTGAGAGGGGGCCACCGCGG	360
QY	361	TCCTCGGCTGTCTGCACCTGTGCGCTGAGCTGCTGAGATCCGGAAGACATGCCAG	420
Db	361	TCCTCGGCTGTCTGCACCTGTGCGCTGAGCTGCTGAGATCCGGAAGACATGCCAG	420
QY	421	TTACCACTGCTTGAATGATAGTGGCTTCTGTTGTCAGTCTCATATAGAACTACAG	480
Db	420	TTACCACTGCTTGAATGATAGTGGCTTCTGTTGTCAGTCTCATATAGAACTACAG	479
QY	481	CTCATCAGGAGGAGATCGGAGCGGTAAAGACACCAACACCATGTTTCGACGAGACT	540
Db	480	CTCATCAGGAGGAGATCGGAGCGGTAAAGACACCAACACCATGTTTCGACGAGACT	539
QY	541	CAAGGATCTCAAGATCAGAGAGTGTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT	600
Db	540	CAAGGATCTCAAGATCAGAGAGTGTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT	599
QY	601	TACAGAGTCTTCAGAGGAGGACGAGGAGCTCAGAGAGCTGCAAGAACCAACCGTGCCAT	660
Db	600	TACAGAGTCTTCAGAGGAGGACGAGGAGCTCAGAGAGCTGCAAGAACCAACCGTGCCAT	659
QY	661	CTCTCAAGACATTCCTTGAGAGAACATACAAAGAAAGTCTTCTTCTTCTTCTTCTTCTTCT	720
Db	660	CTCTCAAGACATTCCTTGAGAGAACATACAAAGAAAGTCTTCTTCTTCTTCTTCTTCTTCT	719
QY	721	GAGCCGAGTCTATCTTACACTTTGCGAGAGAGTATTTGCAAACTGATTTTCCAGAGTT	780
Db	720	GAGCCGAGTCTATCTTACACTTTGCGAGAGAGTATTTGCAAACTGATTTTCCAGAGTT	779
QY	781	TGAACGGCTCAATTTGCACTTCAGAGAACATTTGGCAAGACACAAAATAAAGAAAGACAG	840
Db	780	TGAACGGCTCAATTTGCACTTCAGAGAACATTTGGCAAGACACAAAATAAAGAAAGACAG	839
QY	841	GAATCTTTTGGAAAGAGAGACTTTGAAAACAAATTTGAGAGCAAGAGTTCGACGAGG	900
Db	840	GAATCTTTTGGAAAGAGAGACTTTGAAAACAAATTTGAGAGCAAGAGTTCGACGAGG	895
QY	901	AGTTCCAGTGGAGGTTTATCAAGAAATCTCTTGTGTGAGAGGTTTAAAAATATGTACGA	960
Db	896	AGAGTCCAGTGGAGTTTATCAAGAAATCTCTTGTGTGAGAGGTTTAAAAATATGTACGA	955
QY	961	GGAAGATGAAAACATCCTTGGGCTGTTGGAGGACCCCTTAAAGATTTTAAACAGCTTT	1020
Db	956	GGAAGATGAAAACATCCTTGGGCTGTTGGAGGACCCCTTAAAGATTTTAAACAGCTTT	1014
QY	1021	CAGTACCTTCTGAAACAGAGAGCCATTGCCAAGAGAGGAAAGGGGAGGCTTGA	1080
Db	1015	CAGTACCTTCTGAAACAGAGAGCCATTGCCAAGAGAGGAAAGGGGCA-GCTTGA	1073
QY	1081	GGAGCGCTCATCTATGCTGATAGGAGGAGTATTTCTACATGTTTACTACTTCTT	1140
Db	1074	GGAGCGCTCATCTATGCTGATAGGAGGAGTATTTCTACATGTTTACTACTTCTT	1133
QY	1141	CCCTAAGAGAACACCTCCTGATTTCTCCGCGCATATAAAGCGAGCTCTCACGATTT	1200
Db	1134	CCCTAAGAGAACACCTCCTGATTTCTCCGCGCATATAAAGCGAGCTCTCACGATTT	1193
QY	1201	ATATGAAACGGAGTGGAGTGTCTTAATGCTCCCTGCTTCCATTAATGATTGACGCGA	1260
Db	1194	ATATGAAACGGAGTGGAGTGTCTTAATGCTCCCTGCTTCCATTAATGATTGACGCGA	1253
QY	1261	GTATTGTAATCAGCCCTACTTGTGTTACTCCGTTACATGAAAGACCAAGCAATCCCT	1320
Db	1254	GTATTGTAATCAGCCCTACTTGTGTTACTCCGTTACATGAAAGACCAAGCAATCCCT	1313
QY	1321	GTCCCCCAGCAACCCAGTCTCCTGCTGATTTCCACATCGCTATTTCTGCAAGACATT	1380
Db	1314	GTCCCCCAGCAACCCAGTCTCCTGCTGATTTCCACATCGCTATTTCTGCAAGACATT	1373
QY	1381	TCCATTCCATTTTCATGTTGACAAAGATATGCAATTTGCAATTTGGCAATGGCATCAG	1440
Db	1374	TCCATTCCATTTTCATGTTGACAAAGATATGCAATTTGCAATTTGGCAATGGCATCAG	1433
QY	1441	AAGGCTGATGAAACAGAGAGACTTTCAAGGAAAGCTTAATTTGAAAGATACATTTGAAAT	1500
Db	1434	AAGGCTGATGAAACAGAGAGACTTTCAAGGAAAGCTTAATTTGAAAGATACATTTGAAAT	1490
QY	1501	TCTGACTCCAAAATCAACACAGAGAGTGTAGCGGATCATGACATGTTGTAATGACAGTT	1560
Db	1491	TCTGACTCCAAAATCAACACAGAGAGTGTAGCGGATCATGACATGTTGTAATGACAGTT	1550
QY	1561	TGTTGTACGAGTGGAGGATGGGACAACTCTGTGAAGAAATCTTCAAGGTTTATGACCT	1620
Db	1551	TGTTGTACGAGTGGAGGATGGGACAACTCTGTGAAGAAATCTTCAAGGTTTATGACCT	1610
QY	1621	CAAGGCGCAATGATCTACATTTGAAATCCAGTGCATCTTCTTCTTCTTCTTCTTCTTCT	1680
Db	1611	CAAGGCGCAATGATCTACATTTGAAATCCAGTGCATCTTCTTCTTCTTCTTCTTCTTCT	1670
QY	1681	TGTGGACAGATTAAGAGATTTTACAGAGAGGAGCTTACCTCTCAGACATCCCAATTC	1740
Db	1671	TGTGGACAGATTAAGAGATTTTACAGAGAGGAGCTTACCTCTCAGACATCCCAATTC	1730
QY	1741	CAATGCACTGAGGATGTGCTCTTAATAGGGGAAACAGCCCGAGCTCAAGTGGCCTGAA	1800

```
Db 1731 CAATGCATGAGGATGGTCTTAATAGGGAACAAGCCGAGCTCAAGATGGCTGAA 1790
QY 1801 GAAGAGGCTGGGAAAGCTCAAGGCTTACCTTTGAGCAAGCCACCAGCCCTGGAGGAGA 1860
Db 1791 GAAGAGGCTGGGAAAGCTCAAGGCTTACCTTTGAGCAAGCCACCAGCCCTGGAGGAGA 1850
QY 1861 GAAGAAAAGACAGTAGACCTTCTGTGCTCATATTTCCCTGTGAGGTTGCTCAGCAGCT 1920
Db 1851 GAAGAAAAGACAGTAGACCTTCTGTGCTCATATTTCCCTGTGAGGTTGCTCAGCAGCT 1910
QY 1921 GTGGCAAGGCAAGTTGCAAGCAAGCAAGTTCAAGTATGTCACCATGCTCTTCTCAGA 1980
Db 1911 GTGGCAAGGCAAGTTGCAAGCAAGCAAGTTCAAGTATGTCACCATGCTCTTCTCAGA 1970
QY 1981 CATCGTTGGGTTCACTGCCATCTGCTCCAGTGTCTCAGGTCGACGGTGCAGGTCATCACCATGCT 2040
Db 1971 CATCGTTGGGTTCACTGCCATCTGCTCCAGTGTCTCAGGTCGACGGTGCAGGTCATCACCATGCT 2030
QY 2041 CAATGCATGTACACTCGCTTCGACCAAGAGTGTGGAGAGTGTGATGTCTCAAGAGTGA 2100
Db 2031 CAATGCATGTACACTCGCTTCGACCAAGAGTGTGGAGAGTGTGATGTCTCAAGAGTGA 2090
QY 2101 GACCATTCGCGATGCTATTGTGTAGC--TGGGGGATTACAAAGAGAGTGTATCTCATG 2159
Db 2091 GACCATT--GCGATGCTATTGTGTGGCTTGGGGGATTACAAAGAGAGTGTATCTCATG 2149
QY 2160 CTGTTTCAGATAGCGTGTATGCGCCCTGAAGATGATGGAGTCTCTCATGAAGTTATGTCTC 2219
Db 2150 CTGTTTCAGATAGCGTGTATGCGCCCTGAAGATGATGGAGTCTCTCATGAAGTTATGTCTC 2209
QY 2220 CCATGGAGAACCTATCAAGATGCGAATTTGAGTGTGAGTGTGATGTCTCAAGAGTGA 2279
Db 2210 CCATGGAGAACCTATCAAGATGCGAATTTGAGTGTGAGTGTGATGTCTCAAGAGTGA 2269
QY 2280 TCGTTGGAGTAAATCCCGTACTGCTCTTTTGGAAACAATGCTACTCTGGCTAACA 2339
Db 2270 TCGTTGGAGTAAATCCCGTACTGCTCTTTTGGAAACAATGCTACTCTGGCTAACA 2329
QY 2340 AATTTGAGTCTGCTGAGTGTACCAAGAAAATCAATGTGAGCCCAACAACCTTACAGATTAC 2399
Db 2330 AATTTGAGTCTGCTGAGTGTACCAAGAAAATCAATGTGAGCCCAACAACCTTACAGATTAC 2389
QY 2400 TCAAGACTGTCTGCTGTTGTTTACCCCTCGATCAAGGAGAGTGTCCACCAAACT 2459
Db 2390 TCAAGACTGTCTGCTGTTTACCCCTCGATCAAGGAGAGTGTCCACCAAACT 2449
QY 2460 TCCCTAGTGAATCCCGGAATCTGCCATTTCTGGATGCTTACCAACAAGCAAACT 2519
Db 2450 TCCCTAGTGAATCCCGGAATCTGCCATTTCTGGATGCTTACCAACAAGCAAACT 2509
QY 2520 CAAAACCATGCTTCCAAAAGAAAGATGTGGAAGATGGCAATGCCAATTTTGTAGGCAAG 2579
Db 2510 CAAAACCATGCTTCCAAAAGAAAGATGTGGAAGATGGCAATTTTGTAGGCAAG 2567
QY 2580 CATCAGGAATAGATTAGCAACCTATATACCTATTATTAAGTCTTTGGGTTTGACTCAT 2639
Db 2568 CATCAGGAATAGATTAGCAACCTATATACCTATTATTAAGTCTTTGGGTTTGACTCAT 2627
QY 2640 GAAGATGTGAGAGCTCTGAAGCAGTGTAGGATTTAGATGCTTACACAGCAGTATT 2699
Db 2628 GAAGATGTGAGAGCTCTGAAGCAGTGTAGGATTTAGATGCTTACACAGCAGTATT 2687
QY 2700 AAAATTTGAGGAGCAAGTCAAACTTTTCTCTCTTTTAACTATGACAAATGTACTCAT 2759
Db 2688 AAAATTTGAGGAGCAAGTCAAACTTTTCTCTCTTTTAACTATGACAAATGTACTCAT 2747
QY 2760 TCAGTACTTCAGCTCTTCAAGAAAACCAAAACCTTAAAGAGTACTTTTGTGGAGT 2819
Db 2748 TCAGTACTTCAGCTCTTCAAGAAAACCAAAACCTTAAAGAGTACTTTTGTGGAGT 2807
QY 2820 ATTTCTATTATATACACAGCAGTACTACTCTTACTTCAAAATTCAGCAGCTTTGTACAT 2879
```

```
Db 2808 ATTTCTATTATATATACACAGCAGCTTACTACTCTTACTCAAAATTCAGACCTTTGTACATAT 2867
QY 2880 ATCAGATAATTTAGTCAATTTGTACAAACTGATGAGTCACTGCAATCTCATATCTCTGG 2939
Db 2868 ATCAGATAATTTAGTCAATTTGTACAAACTGATGAGTCACTGCAATCTCATATCTCTGG 2927
QY 2940 TGGATGCGATGCTTATTAAAGTGTGTTGTGATAG--TGTCTCTCAAAAAAATAAAAA 2998
Db 2928 TGGATGCGATGCTTATTAAAGTGTGTTGTGATAGTGTCTCAAAAAAATAAAAA 2987
QY 2999 AAAAAAATAAAAAA 3015
Db 2988 AAAAAAATAAAAAA 3004

RESULT 4
BC028384
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT

BC028384 2657 bp mRNA linear PRI 25-APR-2002
Homo sapiens, Similar to guanylate cyclase 1, soluble, alpha 3,
clone MGC:33150 IMAGE:4792661, mRNA, complete cds.
BC028384
MGC.
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2657)
Strausberg, R.
Direct Submission
Submitted (23-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 46 Row: a Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504212.

FEATURES
source
1..2657
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:33150 IMAGE:4792661"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
388..2460
/codon_start=1
/product="Similar to guanylate cyclase 1, soluble, alpha
3"
/protein_id="AAH28384.1"
/db_xref="GI:20306359"
/translation="MFTKLDKLTGTCFPLLAGQVNPESSEAAAGSSSCKATM
PTCQDIPKNTQESLPQKRTSRVYLHTLAESICKLIFPEFERLNVALQRTLAKHKI
KESKSLREDPEKTIAGQAVAGVPVEIKESLGEVFKICYEEDENLTGVGGTLK
DELNSFTLLKQSSHCQKAGKRGLEDASILCLDEDELHYHYFFPKRTTSLILPGI
IKAAHLYIETEVEVSLMPPCPHNDCEVFNQPLYLYSVHMKTSFSPSPQSSLV
IPTSLFCFKTFPPHFMFKDMTILQFNGIRLNRDRDFOGKPNFEEYFELPKINQT
```


QY	2367	AAATCAATGTCAGCCCAACAACTTACAGATTACTCAAGACTGTCTCGTTTCGGTTTA	2426
Db	2231	AAATCAATGTCAGCCCAACAACTTACAGATTACTCAAGACTGTCTCGTTTCGGTTTA	2290
QY	2427	CCCTCGATCAAGGAGGAACTTCCACCAAACTTCCCTAGTGAATCCCGGAATCTGCC	2486
Db	2291	CCCTCGATCAAGGAGGAACTTCCACCAAACTTCCCTAGTGAATCCCGGAATCTGCC	2350
QY	2487	ATTTTCTGATGCTTTACCAACGAAGCAAACTCAAAACCATGCTTCCAAAAGAAGATG	2546
Db	2351	ATTTTCTGATGCTTTACCAACGAAGCAAACTCAAAACCATGCTTCCAAAAGAAGATG	2410
QY	2547	TGGAAGATGGCAATCCCAATTTTTCAGCAAGCATCAGGAATAGATTAGCAACCTATAT	2606
Db	2411	TGGAAGATGGCAATCCCAATTTTTCAGCAAGCATCAGGAATAGATTAGCAACCTATAT	2470
QY	2607	ACCTATTTATAAGTCTTTGGGCTTTGACTTCATTGAAGATGTAGAGCTCTGAAAGCAC	2666
Db	2471	ACCTATTTATAAGTCTTTGGGCTTTGACTTCATTGAAGATGTAGAGCTCTGAAAGCAC	2530
QY	2667	TTTAGGATTPAGATGGCTAACAGCAGTATTAATAATTTTCAGGAGCCCAAGTCAATCT	2726
Db	2531	TTTAGGATTPAGATGGCTAACAGCAGTATTAATAATTTTCAGGAGCCCAAGTCAATCT	2590
QY	2727	TTCTCCTGTTTAAATGACAAATGACTCACTTCACTTCACTTCACTTCACTTCACTT	2786
Db	2591	TTCTCCTGTTTAAATGACAAATGACTCACTTCACTTCACTTCACTTCACTTCACTT	2650
QY	2787	AAAAAA 2793	
Db	2651	AAAAAA 2657	
RESULT 5			
AK096588			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OCBBF2010420"
/tissue_type="brain"
/clone_lib="OCBBF2"
/dev_stage="fetal"
/note="cloning vector: pME18SFL3"
BASE COUNT 759 a 585 c 640 g 643 t
ORIGIN

Query Match 76.7%; Score 2312.6; DB 9; Length 2627;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2315; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 459 CAGTCTCATATTAAGAACTACAGCTCATCAGGAGGAGATCGCAGCAGGGTAAAGACACCA 518
Db 309 CAATCTCATATAAGAACTACAGCTCATCAGGAGGAGATCAGAGGGTAAAGACACCA 368
QY 519 ACACCATGTTTGCACGAAGCTCAAGGATCTCAAGATCAAGATCAAGAGAGTGTCTTCTCT 578
Db 369 ACACCATGTTTGCACGAAGCTCAAGGATCTCAAGATCAAGATCAAGAGAGTGTCTTCTCT 428
QY 579 TACTGGCACCGTCAAGTTCCTTAACGAGTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAG 638
Db 429 TACTGGCACCGTCAAGTTCCTTAACGAGTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAG 488
QY 639 GCTGCAAGAACACCGTGCCTCATCTGTCAAGACATCTTCTGAGAGAACATACAGAAAGTC 698
Db 489 GCTGCAAGAACACCGTGCCTCATCTGTCAAGACATCTTCTGAGAGAACATACAGAAAGTC 548
QY 699 TTCCTCAAGAAACACCGTGCCTCATCTGTCAAGACATCTTCTGAGAGAACATACAGAAAGTC 758
Db 549 TTCCTCAAGAAACACCGTGCCTCATCTGTCAAGACATCTTCTGAGAGAACATACAGAAAGTC 608
QY 759 GCAAACTGATTTTCCAGAGTTTGAACGGCTTGAACGGCTTGAACGGCTTGAACGGCTTGA 818
Db 609 GCAAACTGATTTTCCAGAGTTTGAACGGCTTGAACGGCTTGAACGGCTTGAACGGCTTGA 668
QY 819 AGCAGCAAAATAAAGAAAGCAGGAAATCTTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 878
Db 669 AGCAGCAAAATAAAGAAAGCAGGAAATCTTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 728
QY 879 CAGAGCAAGCAGTTGCAGCAGGAGTTCAGTGGAGGTTTCAAGAGATCTTCTTGGTGAAG 938
Db 729 CAGAGCAAGCAGTTGCAGCAGGAGTTCAGTGGAGGTTTCAAGAGATCTTCTTGGTGAAG 788
QY 939 AGTTTTAAAAATATGTTACAGAGAGATGAAACATCTTGGGGTGGTGGAGGAGCACC 998
Db 789 AGTTTTAAAAATATGTTACAGAGAGATGAAACATCTTGGGGTGGTGGAGGAGCACC 848
QY 999 TTAAGAGATTTTAAACAGCTTACAGTACCTTCTGAAACAGAGAGAGAGAGAGAGAGAGAG 1058
Db 849 TTAAGAGATTTTAAACAGCTTACAGTACCTTCTGAAACAGAGAGAGAGAGAGAGAGAGAG 908
QY 1059 CAGGAAAAAGGGCAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTT 1118
Db 909 CAGGAAAAAGGGCAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTT 968
QY 1119 TTCTACATGTTTACTTACTTCTTCCCTTAAGAGAACACCTTCCCTTGAATCTTCCCGGCATCA 1178
Db 969 TTCTACATGTTTACTTACTTCTTCCCTTAAGAGAACACCTTCCCTTGAATCTTCCCGGCATCA 1028
QY 1179 TAAAGGCAGCTGCTCAGCTATTATATGAAACGAAGTGAAGTGTCTGTTAAATGCCCTCCCT 1238
Db 1029 TAAAGGCAGCTGCTCAGCTATTATATGAAACGAAGTGAAGTGTCTGTTAAATGCCCTCCCT 1088
QY 1239 GCTTCCATTAATGATTCAGCAGGATTTGTGAATCAGCCCTTACTTGTGTACTCGTTTCA 1298
Db 1089 GCTTCCATTAATGATTCAGCAGGATTTGTGAATCAGCCCTTACTTGTGTACTCGTTTCA 1148
QY 1299 TGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1358
Db 1149 TGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1208

HGEPIKMRIGLHSGSVFAGVGVKMPRYCLFNNVTLANKFESCSVPKINVSPTTYR LKDCPFGVFTPSREELPPNFPSEIPGICHELDAIQQGTNSKPCPKQKDDVEDGNANF LGKASGID*									
BASE COUNT	667	a	500	c	520	g	571	t	
ORIGIN									
Query Match	74.9%; Score 2258; DB 9; Length 2258;								
Best Local Similarity	100.0%; Pred. No. 0;								
Matches 2258; Conservative	0; Mismatches 0; Indels 0; Gaps 0;								
QY	507	TAAGACACCAACACCATGTTCTGCACGAAGCTCAAGGATCTCAAGATCAAGAGAGT	566						
Db	1	TAAGACACCAACACCATGTTCTGCACGAAGCTCAAGGATCTCAAGATCAAGAGAGT	60						
QY	567	GTCCCTTTCCTTACTGGCACCAGGTCAAGTTCCTTAACGAGTCTTCAGAGGAGCAGAG	626						
Db	61	GTCCCTTTCCTTACTGGCACCAGGTCAAGTTCCTTAACGAGTCTTCAGAGGAGCAGAG	120						
QY	627	GAAGCTCAGAGAGCTGCAAGCAACCGTCCCATCTGTCAAGACATTCCTGTGAAGAACA	686						
Db	121	GAAGCTCAGAGAGCTGCAAGCAACCGTCCCATCTGTCAAGACATTCCTGTGAAGAACA	180						
QY	687	TACAAGAAAGTCTTCTCAAGAAACACGAGCGGAGCGAGTCTATCTTCACACTTTGG	746						
Db	181	TACAAGAAAGTCTTCTCAAGAAACACGAGCGGAGCGAGTCTATCTTCACACTTTGG	240						
QY	747	CAGAGAGTATTGCAAACTGATTTTCCAGAGTGTGAAGCGGTGAATGTGCACCTTCAGA	806						
Db	241	CAGAGAGTATTGCAAACTGATTTTCCAGAGTGTGAAGCGGTGAATGTGCACCTTCAGA	300						
QY	807	GAACATTGCAAGCAACAAATAAAGAAAGACAGGAATCTTTGGAAAGAGAGACTTTG	866						
Db	301	GAACATTGCAAGCAACAAATAAAGAAAGACAGGAATCTTTGGAAAGAGAGACTTTG	360						
QY	867	AAAAACAATTGCAGACCAAGCAGTTGCAGCAGGAGTTCAGTGGAGGTTATCAAGAAT	926						
Db	361	AAAAACAATTGCAGACCAAGCAGTTGCAGCAGGAGTTCAGTGGAGGTTATCAAGAAT	420						
QY	927	CTCTTGGTGAAGAGTTTTTAAATATGTTACGAGGAGATGAAACATCTCTTGGGGTG	986						
Db	421	CTCTTGGTGAAGAGTTTTTAAATATGTTACGAGGAGATGAAACATCTCTTGGGGTG	480						
QY	987	TTGGAGGACCTTAAAGATTTTTTAAACAGCTTCAGTACCTCTTGAACAGAGAGCC	1046						
Db	481	TTGGAGGACCTTAAAGATTTTTTAAACAGCTTCAGTACCTCTTGAACAGAGAGCC	540						
QY	1047	ATTGCCAAGAAGCAGGAAAAAGGGCAGGCTTGAGGACGCTCCATCTATGCTGGATA	1106						
Db	541	ATTGCCAAGAAGCAGGAAAAAGGGCAGGCTTGAGGACGCTCCATCTATGCTGGATA	600						
QY	1107	AGGAGGATGATTTCTACATGTTTACTTCTTCCCTAAGAGACCCCTCCCTGATTC	1166						
Db	601	AGGAGGATGATTTCTACATGTTTACTTCTTCCCTAAGAGACCCCTCCCTGATTC	660						
QY	1167	TTCCCGGCATCAAAAGCAGCTGCTCAGCTATATATCAAGCGGAGTGGAGTCTCT	1226						
Db	661	TTCCCGGCATCAAAAGCAGCTGCTCAGCTATATATCAAGCGGAGTGGAGTCTCT	720						
QY	1227	TAATGCTCCCTGCTTCCATATGATTGACGGAGTTTGTGAATCAGCCCTACTTGTGT	1286						
Db	721	TAATGCTCCCTGCTTCCATATGATTGACGGAGTTTGTGAATCAGCCCTACTTGTGT	780						
QY	1287	ACTCCGTTACATGAAGACACCAAGCCATCCCTGTCCCGCAGCAACCCAGTCTCCG	1346						
Db	781	ACTCCGTTACATGAAGACACCAAGCCATCCCTGTCCCGCAGCAACCCAGTCTCCG	840						
QY	1347	TGCTGATCCCATCGCTATCTGCAAGACATTTCCATTCATTTTCATGTTTGACAAAG	1406						
Db	841	TGCTGATCCCATCGCTATCTGCAAGACATTTCCATTCATTTTCATGTTTGACAAAG	900						
QY	1407	ATATGACAATTTGCAATTTGGCAATGGCATCAGAAGGCTGATGAACAGGAGACTTTC	1466						
Db	901	ATATGACAATTTGCAATTTGGCAATGGCATCAGAAGGCTGATGAACAGGAGACTTTC	960						

QY	1467	AAGGAAAGCCTAATTTTGAAGATACTTTGAAATCTGACTCCAAAAATCAACCAGACGT	1526						
Db	961	AAGGAAAGCCTAATTTTGAAGATACTTTGAAATCTGACTCCAAAAATCAACCAGACGT	1020						
QY	1527	TTAGCGGGATCATGACTATGTTGAATATGCAGTTTGTGTACGAGTGAGGAGATGGACA	1586						
Db	1021	TTAGCGGGATCATGACTATGTTGAATATGCAGTTTGTGTACGAGTGAGGAGATGGACA	1080						
QY	1587	ACTCTGTGAAGAAATCTTCAAGGGTTATGACCTCAAGGCCAAATGATCTACATTTGTG	1646						
Db	1081	ACTCTGTGAAGAAATCTTCAAGGGTTATGACCTCAAGGCCAAATGATCTACATTTGTG	1140						
QY	1647	AATCCAGTGCATCTTTGTTTGGGTCACTTGTGTGGACAGATTAGAAGATTTTACAG	1706						
Db	1141	AATCCAGTGCATCTTTGTTTGGGTCACTTGTGTGGACAGATTAGAAGATTTTACAG	1200						
QY	1707	GACGAGGCTCTACCTCTCAGACATCCCAATTCACAATCACAATGAGGATGTGCTTAA	1766						
Db	1201	GACGAGGCTCTACCTCTCAGACATCCCAATTCACAATCACAATGAGGATGTGCTTAA	1360						
QY	1767	TAGGGAACAAGCCGAGCTCAAGATGGCTGAAGAGAGGCTGGGGAAGCTCAAGGCTA	1826						
Db	1261	TAGGGAACAAGCCGAGCTCAAGATGGCTGAAGAGAGGCTGGGGAAGCTCAAGGCTA	1320						
QY	1827	CCCTTGAGCAAGCCCAAGCCCTGGAGAGGAGAGAAAGAACAGTAGACCTTCTGT	1886						
Db	1321	CCCTTGAGCAAGCCCAAGCCCTGGAGAGGAGAGAAAGAACAGTAGACCTTCTGT	1380						
QY	1887	GCTCCATATTTCCCTGTGAGTTGCTCAGCAGCTGTGCGAAGGCCAAGTTGCAAGCCA	1946						
Db	1381	GCTCCATATTTCCCTGTGAGTTGCTCAGCAGCTGTGCGAAGGCCAAGTTGCAAGCCA	1440						
QY	1947	AGAAGTTCAATATGTCACCATCTCTTCAACATCGTTGGTTCACCTGCCATCTGCT	2006						
Db	1441	AGAAGTTCAATATGTCACCATCTCTTCAACATCGTTGGTTCACCTGCCATCTGCT	1500						
QY	2007	CCAGTGTCAACGCTGACAGTCAATCACCATGCTCAATGCATGTACACTCGCTTCGACC	2066						
Db	1501	CCAGTGTCAACGCTGACAGTCAATCACCATGCTCAATGCATGTACACTCGCTTCGACC	1560						
QY	2067	AGCAGTGTGGAGAGCTGGATGCTCAAGGTGGAGACCATTTGGCGATGCTATTTGTAG	2126						
Db	1561	AGCAGTGTGGAGAGCTGGATGCTCAAGGTGGAGACCATTTGGCGATGCTATTTGTAG	1620						
QY	2127	CTGGGGATTACAAAGAGAGTACTACTCTGTCAGATAGCGCTGATGCGCCTGA	2186						
Db	1621	CTGGGGATTACAAAGAGAGTACTACTCTGTCAGATAGCGCTGATGCGCCTGA	1680						
QY	2187	AGATGATGAGCTCTCTGATGAAGTTATGTCCTCCCATGGAGAACCTATCAAGATCGAA	2246						
Db	1681	AGATGATGAGCTCTCTGATGAAGTTATGTCCTCCCATGGAGAACCTATCAAGATCGAA	1740						
QY	2247	TTGGACTGCACCTCTGATCAGTTTTCCTGGCGCTGCTGGAGTTAAATGCCCGTACT	2306						
Db	1741	TTGGACTGCACCTCTGATCAGTTTTCCTGGCGCTGCTGGAGTTAAATGCCCGTACT	1800						
QY	2307	GTCTTTTGGAAACAATGTCACTCTGGCTTAACAAATTTAGTCTCTGCAGTGTACACGAA	2366						
Db	1801	GTCTTTTGGAAACAATGTCACTCTGGCTTAACAAATTTAGTCTCTGCAGTGTACACGAA	1860						
QY	2367	AAATCAATGTACGCCCCAACACTTACAGATTACTCAAGACTGCTCTGTTTTCGTGTTA	2426						
Db	1861	AAATCAATGTACGCCCCAACACTTACAGATTACTCAAGACTGCTCTGTTTTCGTGTTA	1920						
QY	2427	CCCTCTGATCAGGGAGAACTTCCACCAACTTCCCTAGTAAATCCCGGATCTGCC	2486						
Db	1921	CCCTCTGATCAGGGAGAACTTCCACCAACTTCCCTAGTAAATCCCGGATCTGCC	1980						
QY	2487	ATTTTCTGATGCTTACCAACAGGAAACAACTCAAAACCATGCTTCCAAAAGAAAGATG	2546						
Db	1981	ATTTTCTGATGCTTACCAACAGGAAACAACTCAAAACCATGCTTCCAAAAGAAAGATG	2040						

```

QY 2547 TGGAGATGCAATGCCAATTTTGGCAAGCATCAGGAATAGATTAGCAACCTATAT 2606
Db 2041 TGGAGATGCAATGCCAATTTTGGCAAGCATCAGGAATAGATTAGCAACCTATAT 2100
QY 2607 ACCTATTATAAGTCTTTGGGTTGACTCATTTGAAGATGTGAGAGCTCTGAAGCAC 2666
Db 2101 ACCTATTATAAGTCTTTGGGTTGACTCATTTGAAGATGTGAGAGCTCTGAAGCAC 2160
QY 2667 TTTAGGATGTAGATGGCTTAACAAGCAGTATTAAATTTTTCAGGAGCAAGTCAACATCT 2726
Db 2161 TTTAGGATGTAGATGGCTTAACAAGCAGTATTAAATTTTTCAGGAGCAAGTCAACATCT 2220
QY 2727 TTTCTCTGTTTAACATGACAAATGTACTCACTTCACT 2764
Db 2221 TTTCTCTGTTTAACATGACAAATGTACTCACTTCACT 2258

RESULT 7
BTGUCY73
LOCUS 3261 bp mRNA linear MAM 12-SEP-1993
DEFINITION Bovine mRNA for large subunit of guanylate cyclase (EC 4.6.1.2).
ACCESSION X54014
VERSION X54014.1 GI:405
KEYWORDS guanylate cyclase.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 3261)
Koesling,D.
Direct Submission
Submitted (04-OCT-1990)
2 (bases 1 to 3261)
Koesling,D., Harteneck,C., Humbert,P., Bosserhoff,A., Frank,R.,
Schultz,G. and Bohme,E.
The primary structure of the larger subunit of soluble guanylyl
cyclase from bovine lung. Homology between the two subunits of the
enzyme
FEBS Lett. 266 (1-2), 128-132 (1990)
90306336
PUBMED
1973124
FEATURES
Location/Qualifiers
1..3261
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="adrenal medulla"
/clone_lib="lambda Zap"
291..2366
/note="guanylate cyclase large subunit"
/codon_start=1
/protein_id="CAA37960.1"
/db_xref="GI:406"
/db_xref="SWISS-PROT:P19687"
/translation="MFCALIKLQITGDPCFLLAPGOVPEPLGEATGSGPASTPGQ
PGVCPGVPDKNPGRPLRRNTSRVLTILAESICKLIPFEFERLNLAQRTLAHK
IKENKSLREDEKIVDQAIAGVPVEIKESLGEELFKIYEDEYILGVVGTI
KDFLNSFTLLKQSSHCQAEKKREFEDASILCLDKDPDLYVYFFPKRITSLILPG
IIKAAARILYETEVEVSTSPREHQDREFVQPCELYSVHRSRPHPPGKPVSSL
VPASLFCKTFPEHMLDRDMSILOLGHGIRRLMSRDVQKPFDEYFELTPKLSQ
TFSGIITMLNQFLVVRWDSNKKSSRVNDLKGOMIYVESSSILFLGSPVDRL
DFTGRGLISDIPFHNALRDVLIGQARQDGLKRLKATLQAOHAEERK
TVDLLCSIFPEFARLMQGHVQAKRFGNVTMLFSDIVGTIACSQCSPLOVITMLN
ALYTRFDQCGELDVYKVTIGDAYCVAGLGHKSDTHAVQIALMKMLKMEHSVYS
PHGEIKMRIGLSGVSFVAGVGVGMKPRVCLFGNVTLANQIFESCSVPKINVSPTTY
RLKDKCGFVFTSRRELPNPFPSDIPGICHELAYQQGTTTSKPFQKDVSEANAN
FLKASGID"
BASE COUNT 845 a 789 c 814 g 813 t
ORIGIN
Query Match 53.8%; Score 1620.6; DB 4; Length 3261;
Best Local Similarity 79.8%; Pred. No. 0;
Matches 2006; Conservative 0; Mismatches 479; Indels 28; Gaps 7;

```

```

QY 511 AGACACCAACACCATGTTCTGCACGAGCTCAAGGATCTCAAGATCAGAGGAGTGTC 570
Db 278 AGGAGCAATCACCATGTTTTCGCGAAGCTGAAGACCTGCAGATCAGGGGACTGCC 337
QY 571 TTTCTCTCTACTTGCACCAAGTCAAGTTCCTAACGAGTCTTTCAGAGGAGGAGGAG 630
Db 338 TTTCTCTCTCTCTGCGCCCGGCGAGGTCTCCGACAGAGCCCTTGGGGGAAGCCAC 397
QY 631 CTCAGAGAGCTGCAAAAGCAACCGTGCC--CATCTGTCAAGACATCTTCTCAGAG 687
Db 398 TGGGCCAGCATCGACCCAGGCGAGCCAGCGGTCTGTCCAGGTGTCCCCACAGAAC 457
QY 688 ACAAGAAAGTCTTCTCTCAAAAGAAACCACTCGGAGCGGAGTCTATCTTTCACAC 747
Db 458 TCCCGGAAGGCTTCCCGCGAGGAGACACCGCGCGGTCTATCTTTCACACTTTGC 517
QY 748 AGAGAGTATTGTCAAACTGATTTTCCAGAGTTTGAAGGCTGAATGTTTCACACTC 807
Db 518 AGAGAGTATTGTCAAACTGATTTTCCAGAGTTTGAAGGCTGTGAACCTTCCCTTC 577
QY 808 AACATTGGCAAGACACAAATAAAGAAAGCAGGAATCTTTGGAAAGAGAGACTTTGA 867
Db 578 AACATTGGCAAGACACAAATAAAGAAAGCAGGAATCTTTGGAAAGAGAGACTTTGA 637
QY 868 AAAAACAATTTGCAGAGCAAGCAGTTCGAGCAGGAGTTCCAGTGAGGTTTATCAAG 927
Db 638 AAAAATAGTTGTGACCAAGCAATTCGAGCAGGAGTTCCAGTGAGGATCATCAAGA 697
QY 928 TCTTGTGTAAGAGGTTTAAATATGTTTACGAGGAGGATGAACATCTTGGGGTGT 987
Db 698 TCTTGTGTAAGAGGTTTAAATATGTTTAAATATGTTTAAATATGTTTAAATATG 757
QY 988 TGGAGGCAACCTTAAAGATTTTAAACAGCTTCAGTACCTTCTGAAACAGAGCAG 1047
Db 758 TGGAGGCAACCTTAAAGATTTTAAATATGTTTAAATATGTTTAAATATGTTTAA 817
QY 1048 TTCCCAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1107
Db 818 CTGCTCAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 877
QY 1108 GGAGGATGATTTTCTACATGTTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1167
Db 878 AGATCCCGATGTTTATATGTTTACTTATTTTCTTCTTCTTCTTCTTCTTCTTCT 937
QY 1168 TCCCGGATCATTAAGGAGCAGTCTCAGCTATTATATGAAACGAGGAGTGAAGTGT 1227
Db 938 CCCCAGGATCATTAAGGAGCAGTCTCAGCTATTATGAGACGAGGAGTGAAGTGT 997
QY 1228 AATGCTCTCTGCTTCCATTAATGATTGACGAGGAGTGTGAAATCAGCCCTTCT 1287
Db 998 GAGCGCTTCTGCTTCCATTAATGATTGACGAGGAGTGTGAAATCAGCCCTTCT 1057
QY 1288 CTCGGTTCATCAAAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1347
Db 1058 CTCGGTTCATCAAAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1117
QY 1348 GGTGATTTCCACATCGCTATTCTGCAAGACATTTTCAATTCATGTTTTCAGCA 1407
Db 1118 GGTGATTTCCACATCGCTATTCTGCAAGACATTTTCAATTCATGTTTTCAGCA 1177
QY 1408 TATGACAAATTTCTCAATTTGGCAATGGCATGAGAGGCTGATGAACAGGAGAGCT 1467
Db 1178 TATGAGCATCTCGAGCTCGGCCACGCGATCAGCGGCTGATGAGCAGGAGAGAG 1237
QY 1468 AGGAAAGCCTTAATTTTGAAGAAATACATTTGAAATTCGACTCCAAATCAACAG 1527
Db 1238 AGGAAAGCCTTAATTTTGAAGAAATACATTTGAAATTCGACTCCAAATCAACAG 1297
QY 1528 TAGCGGATCATGACTATGTTGAATATGCAAGTGTGTTGTTGACGAGTGGAGAG 1587
Db 1298 CAGCGGATCATGACTATGTTGAATATGCAAGTGTGTTGTTGACGAGTGGAGAG 1357

```


FLANSFTLLKQSHHQEABRERLRDASLCLDKDQDFLVNYVFFPRRTALLPLLGI
KAARILYFHVHVMWPCKSCTEFVQNYLILYVHVASTKPSLSPGKPOSSVILI
PTSLFCITFHFMLRDALALQNGRNVKLYRNNKDFGKFNFEFFELTPKINOTIF
SGIWMNMOFVIRVRNNDVKKSRQADLKGOMIVESSALVGLSPCVRLEDV
TGRGLYDITPHNALRDVYLGEQARQDLKRLKATLHAHQALFEERKTI
DLIACITFPEVAQLOQWOGIOVAKFNFTMLSDIVGFTAICOSPCLAVITMALAL
YTRFDQOCQDVKYETIGDAYCVAGHGRHSRDTVAQIALKCMKLWELSNVMSPH
GEPKWRIGLHGSFVAGVGVKMPKIPGLFNGNVTLOHNPESCVPRKINVPSTYRL
KDCQCPYFTRSRLEPNFSDIPCLGHFADVAHQGNPSKPEKDAQDGNANF

LGRASGVD			
BASE COUNT	1340 a	1069 c	1042 g
ORIGIN	1340 t	1042 g	1324 t

Query Match	52.28;	Score 1573.6;	DB 10;	Length 4775;
Best Local Similarity	76.88;	Pred. No. 0;		
Matches 2032;	Conservative	0;	Mismatches 589;	Indels 25;
				Gaps 8;

374	QY	CTCGACCCGTGTCGCCCTGAGCTGCCTGACAGTGACAAATGCCAGTATCCAGTGTACCTGTCCT	433
314	Db		373
314	Db	CTCGACGGATGCCCTTGCCCTGCCGTGCCGTATCATGATGCTGGGCCACTCGTGTCTTT	373
434	QY	TGAATTGATAGTGGCTTCTGCTTTGTCAGTCTCATATAAGAACTACAGCTCATCAGGAGGA	493
374	Db		430
494	QY	GATCGCAGCAGGGTAAAGAGACACCAACCATGTTCTGCACGAAGCTCAAGGATCTCAAG	553
431	Db		486
431	Db	AACTCACTGC ---CAAGCTCCAGGAACACCATGTTCTGCGAGGAAGTTCAAAGATCTCAAG	486
554	QY	ATCACAGGAGAGTGTCTTTCTCTTACTGGCACAGGTCAGTTGCTTAAGGAGTCTTCA	613
487	Db		546
487	Db	ATCACGGGGAGTGTCTTTCTCTTACTGGCTCTGGTCAAGTTCCTACGGAGCCAATA	546
614	QY	GAGGAGCAGCAGGAAGCTCAGAGAGCTGCAAAAGCAACCGTGCCCACTCTCTCAAGACATT	673
547	Db		606
547	Db	GAGAGGTGCCCGAGTCTCTGAGAGCTGCCAGGCCACTCTGCCCACTGCCAGGAGTTT	606
674	QY	CCTGAGAAGAACATACAAGAAAGTCTTCTCTCAAGAAAAACCAAGTCGGAGCCGAGTCTAT	733
607	Db		663
607	Db	GCTG ---AGAATTGCAGAGGGAGTCAACCCCAAGAAAGACAAAGCCCAACAGAGTCTAC	663
734	QY	CTTCACACTTTGGCAGAGAGTATTTGCCAACTCATTTTCCACAGTTTGAACGCCTGAAT	793
664	Db		723
664	Db	CTACACCCCTGGCAGAGAGTATTTGGCAAGCTCATTTTCCACAGTTTGAAGACTTGAAAC	723
794	QY	GTTCGCACTTCAGAGAACATTGGCAAGACAAAAATAAAGAAAAAGCAGGAAATCTTTGGAA	853
724	Db		783
724	Db	CTTTCGACTTCAGAGACATTGGCAAGACCAAAATAAAGGAAAAACAGGAATTTCTTCAGAA	783
854	QY	AGAGAAGACTTTGAAAAACAATTTGCAGAGCAAGCAGTTGCGACGAGAGTTTCAGTGGAG	913
784	Db		843
784	Db	AAAGAAGACTTTGAAAGAAATAATTGCAGAGAAGACAAATTGCGCAGGTGTCCAGTGGAG	843
914	QY	GTATTACAAAGAAATCTCTTGCTGAAGAGTTTAAAAATGTTTACGAGGAAGATGAAAC	973
844	Db		903
844	Db	GTGCTCAAGACACTCTGGGTGAGGAAGTCTTCAAGATCTGTTATGAGGAAGATGAGCAC	903
974	QY	ATCCTTTGGGTGTTGGAGCACCCCTTAAAGATTTTTTAAACAGCTTCAGTACCCCTCTG	1033
904	Db		963
904	Db	ATCCTGGCGTGTGTGGCGCACCCCTGAAGGACTTTTCTCAACAGCTTTCAGCAGCTCCTG	963
1034	QY	AAACAGCAGCCATTGTTGCCAAGAAGCAGGNAAGGGGACGGCTTGAGGACGGCTCCATT	1093
964	Db		1023
964	Db	AAGCAGAGCAGCCACTGTTGCCAGGAGCGGAGAGCGGGGGCGACTGGAAGATGCTCCATC	1023
1094	QY	CTATGCTTGGATTAAGGAGGATGATTTTCTACATGTTTACTTACTTCTTCCCTTAAGAGAAC	1153
1024	Db		1083
1024	Db	TTTGCTTGGCAAGCAGCCAGGACTTTCTAAATGTTTACTTACTTCTTCCCGAAGAGAAC	1083
1154	QY	ACCTCCGCTGATCTTCCCGGCATCATAAAGGCAGCTGCTCACGTATATATGAACGGAA	1213
1084	Db		1143
1084	Db	ACAGCCTGTCTTCCCGGTATCATTAAGCGGCTGCTCGCATCTACTGTACGAAGGCAC	1143

Qy	1214	GTGAAGTGTCTGTTAATAGCCTCCCTGCTCCATTAATGATGTGACGAGTTGTGTAATCAG	1273
Db	1144	GTGGAGTGTCCCTGTATGCTCCTGCTCCGAAAGTGTACCGAATTCGTGTGAACCG	1203
Qy	1274	CCCTACTTGTGTACATCCGTTTCACATGAAAGACCAAGCCATCCCTGTCCCCCAGCAA	1333
Db	1204	CCCTATTGTCTTACTCTGGTCCAGTGAAGAGACCAAGCCTCTCTGTCCCCAGGCAA	1263
Qy	1334	CCCAGTCCCTCGTGTGTATTTCCACATTCGTAATCTTCAAGACATTTCCATTTCCATTTTC	1393
Db	1264	CCCAGTCCCTCGTGTGTATTTCCACATTCGTAATCTTCAAGACATTTCCATTTCCATTTTC	1323
Qy	1394	ATGTTTGACAAAAGATATGACAAATCTGCAATTTGGCAATGGCATCAGAAGCTGTATGAAC	1453
Db	1324	ATGCTGGACCGTGACCTGGCCATTCTGCAGCTGGGTAAATGGCATCAGAAGCTGTGTGAAC	1383
Qy	1454	AGGAGACATTTCAAGGAAGCCCTAATTTTGAAGAATACTTTGAATTTCTGACTCCAAA	1513
Db	1384	AAGAGGACATTTCCAAGGAAGCCCAATTTTGAAGAGTTCTTTTGAATTTCTTAATCTCCAAA	1443
Qy	1514	ATCAACCAGACGTTTACGGGGATCATGACTATCTTGAATATGACATTTGTGTACGAGTG	1573
Db	1444	ATCACACGACATTTAGCGGCATCATGACATTTTGAATATGCGATTTGTATACGGGTG	1503
Qy	1574	AGGAGATGGGCAACCTCTGTGAAGAAATCTTCAAGGTTTATGGACCTCAAGGCCAAATG	1633
Db	1504	AGGAGATGGGATTAATTTAGTGAAGAAATCTTCAAGGTTTATGGACCTCAAGGTCAAATG	1563
Qy	1634	ATCTACATTTGTAATCCAGTGCATTTGTTTTTGGGGTCACCCCTGTGTGGACATTTA	1693
Db	1564	ATCTACATCTGTTGAATCCAGTGCATTTGTTTTTGGGGTCACCCCTGTGTGGACATTTG	1623
Qy	1694	GAAGATTTTACAGGACGAGGGCTTACCTCTCAGACATCCCAATTCACATGACATGAGG	1753
Db	1624	GAAGATTTACAGGACGAGGGCTTCTATCTCTTGACATCCGATTCATATGCCCCGTGAGG	1683
Qy	1754	GATGTGCTCTTAATAGGGGAACAGCCGAGCTCAAGATGGCCTGAAGAGAGGCTGGGG	1813
Db	1684	GATGTTGCTCTGATAGGGGAACAGGCACGGCTCAAGATGGCCTCAAGAGAGGTTGGGA	1743
Qy	1814	AAGCTGAAGGCTTACCCTTGAGCAAGCCCAAGCCCTGGAGGAGGAGAAGAAAAGACA	1873
Db	1744	AAGCTGAAGCAACCTTGGAGCATGCCACCAAGCCCTGGAGGAGGAGAAGAAAGACG	1803
Qy	1874	GTAGACTTCTGTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAGCTGTGGCAAGGCCAA	1933
Db	1804	GTAGATCTCTGTGTCTTATCTTTCCCTTGAGGTTGCTCAGCAGCTGTGGCAAGGACAA	1863
Qy	1934	GTGTGTCAAGCCCAAGATTCAGTATGTCAACCATGCTCTTCTCAGACATCGTGTGGTTTC	1993
Db	1864	ATTGTCAAGCCCAAGATTCAGTATGAGTCAACCATGCTCTTCTCAGATTCGTAGGGTTTC	1923
Qy	1994	ACTGCCATCTGCTCCAGTCTCACCGCTCAGGTCATCACCATGCTCAATGCTACGTGTAC	2053
Db	1924	ACTGCCATCTGCTCTCAGTGTTCACCTCTGCAAGTCAATCACCATGCTCAACGCTCTAT	1983
Qy	2054	ACTCGCTTCACACAGAGTGTGGAGAGCTGGATGCTTACAAGGTGGAGACCATTTGGCGAT	2113
Db	1984	ACTCGCTTTGACCCACAGTGTGGAGAGCTGGATGCTTACAAGGTGGAGACCATTCGGAGAT	2043
Qy	2114	GCCTATTGTGTAGCTGGGGATTACAAAGAGAGTGTACTCATGCTGCTTCAGATAGCG	2173
Db	2044	GCATATTGTGTGCGAGGGGATTGCACAGAGAGTGTACACATGCTGTCTCCAGATAGCA	2103
Qy	2174	CTGATGGCCCTGAAGATGATGGAGCTCTCTGATGAAGTTTATGTCTCCCCATGGAGAACCT	2233
Db	2104	CTGATGGCCCTGAAGATGATGGAGCTCTCCAATGAGGTCTATGTCTCCCCACGGAGAACCT	2163
Qy	2234	ATCAAGATGCGAATTTGGACTGCACCTCTGGATCAGTTTTTGTCTGGCGCTGTGTGGATTAA	2293
Db	2164	ATCAAGATGCGAATTCGAGCTGCATTTCTGGGTCAAGTGTGTCTGGAGAGTTGTCTGGAGTGAAG	2223
Qy	2294	ATCCCCCGTTACTGTCTCTTTTGGAAAACATGTCACTCTGGCTAACAAATTTGATGCTCTGC	2353

||||| 2224 ATGCCCGATATTCCTGTTGGAAACAATGTCACCTCTGGCTAACAAATTTGAGTCTGC 2283
Db
||||| 2354 AGTGTACCAACGAAATCAATGTACGCCCAACAACTTACAGATTACTCAAGACTGTCT 2413
QY
||||| 2284 AGTGTGCTCGGAATCAATGTACGCCCAACAACTTACAGATTACTCAAGACTGTCT 2343
Db
||||| 2414 GGTTCGTTTACCCCTCGATCAAGGAGGAACTTCCACCAAACTTCCCTAGTGAATC 2473
QY
||||| 2344 GGTTCGTTTACCCCTCGATCAAGGAGGAACTTCCACCAAACTTCCCTAGTGAATC 2403
Db
||||| 2474 CCGGAATCTGCCATTTCTGGATGCTTACCAA---CAAGGACAACTCAAAACCATGC 2530
QY
||||| 2404 CTTGGAATCTGTCACCTTCTGGATGCTTACCAATCAAGGACCTAATTTCAACCATGG 2463
Db
||||| 2531 TTCCAAAAGAGATGGAAGATGCAATGCCAATTTTATAGGCAAGCATCAGGAATA 2590
QY
||||| 2464 TTCCAGCAGAAAGATGCAAGATGGAAGCATCTTCTTAGGCAAGCATCAGGGTA 2523
Db
||||| 2591 GATTAGCAACCTATATACCTATTATAGTCTTTGGGGTTTGAATTTGAAGATGTGA 2650
QY
||||| 2524 GATTAGTACGCGCATGTTCTGTGTTGATGCTTTGAGGGGTGCTCAGAATCTCTGAA 2583
Db
||||| 2651 GAGCCTCTGAAGCACTTTAGGATGTTAGTGGCTAACAAAGCAGTATTAAATTTGAG 2710
QY
||||| 2584 ATGACTTTAGGATGCAAGATGATAAACTGTGTTAAATTTGAGGGGTGAACCCAGT 2643
Db
||||| 2711 AGCCAAGTCAACATCTTCTCTGTTTAAACATGACAAATGTACTCCTTCACTTCA 2770
QY
||||| 2644 TTTCTCTGTTTCCATTTAAGATGGAAGAAAGATTTCAATGCTTCAGCGCTTCA 2703
Db
||||| 2771 GC-TCTTCAAGAAAGAAAGAAAGCTTAAAGTACTTTTG---TGGAGTATTCTA 2826
QY
||||| 2704 CTTCTCTCTTATTAAAGAAAGAAAGCTTAAAGTACTTTTG---TGGAGTATTCTA 2763
Db
||||| 2827 TTATATAACGACGACTTACTACCTGACTCAAAATTCAGACCTTGTACATATATCAGAT 2886
QY
||||| 2764 TTTGTTTAGAGGAGTCACTTCTGCTGCAATTCACACAGTGTATATTCAGAC 2823
Db
||||| 2887 AATTGAGTCAATGTGA-----CAAACTGATGAGTCACTGCAATCTCATATCTCCGGTG 2941
QY
||||| 2824 AATTGAGTCAATGTGA-----CAAACTGATGAGTCACTGCAATCTCATATCTCCGGTG 2883
Db
||||| 2942 GAATGCCATGTTTAAAGTGTGTTGATGAGTCTGCTCAAAAGAAAGAAAGAAAGAA 3001
QY
||||| 2884 AACTG---TGGTTATTACAGTGTGCTGGTGTGTTGCTGTAATCTCTGAATAGAA 2940
Db
||||| 3002 AAAAAA 3007
QY
||||| 2941 ATGACA 2946
Db

RESULT 9
AF297082
LOCUS
DEFINITION
Mus musculus soluble guanylyl cyclase alpha 1 subunit mRNA,
complete cds.
AF297082
VERSION
AF297082.1 GI:10442713
KEYWORDS
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2445)
AUTHORS
Sharina, I.G., Krumenacker, J.S., Martin, E. and Murad, F.
TITLE
Genomic organization of alpha 1 and beta 1 subunits of the mammalian
soluble guanylyl cyclase genes
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (20), 10878-10883 (2000)
MEDLINE
20461464
PUBMED
10984516
REFERENCE
2 (bases 1 to 2445)
AUTHORS
Krumenacker, J.S., Sharina, I.G., Martin, E.S. and Murad, F.

Direct Submission
Submitted (17-AUG-2000) Integrative Biology and Pharmacology,
University of Texas Houston Medical School, 6431 Fannin St.,
Houston, TX 77030, USA
Location/Qualifiers
1..2445
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
286..2361
/codon_start=1
/product="soluble guanylyl cyclase alpha 1 subunit"
/protein_id="Agi1446.1"
/db_xref="GI:10442714"
/translation="MFCRKKDKITGECFLLAPGVQPKETPEVAGSGCQATL
PICOVPEKNAEGLPQKTSRNVVHTLALSIKLIPECEERLALQRTLAKHKI
ENRKSSEKDEKIIAERAIAGAPVEALKDSIGLEFKICYEEDEHILGVVGTGLK
DELNESPILLKQSSQCEAERGRLEDSILCKDQDFLNVYFFPKRTYALLLPGI
IKAAKILYSHVESVLMPPSCDTEFNQPIYLLSVHVKSTKPSLSPQPSLV
IPASLFTFPFHEMLDRDLAILQLNGIRLVNKRDFQGNPFEEFFILPKINOT
FSGIMTLMNQFIVRVRWDSVYKSSRYMDLKGOMIYVESAILFLGSPCVRLD
FTGRGLYLSDIPIHNAIRDVLIGEORADQGLKRLGLKATLEHAHOLEBEKRT
VLLCSIPPEVAQOLWGOIWOAKFSEVTLFSDIVGTATCSOSPLOVITLNA
LYTRDQCGELDYKVEITGDAYCVAGSLHRESRTHAVQIALMALKMELSNEYMS
HGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNNVTLANKFESCVPKINVSPTIR
LKDKCPGFVTPRSREELPPFPDIPGICFLDAYHHQGNPKWPFQDKVEDGNAN
FLGKASGVD"
BASE COUNT 635 a 624 c 639 g 547 t
ORIGIN
Query Match 51.5%; Score 1552; DB 10; Length 2445;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 1810; Conservative 0; Mismatches 380; Indels 10; Gaps 2;
QY 420 GTTACCAGTCTCTTCAATTTGATGCTGCTTTGTCAGTCTCATATAAGAACTACA 479
Db 189 GTCACCTCTCTCTTCAATTTGATGCTGCTTTGTCAGTCTCATATAAGAACTACA 248
QY 480 GTCATCAGAGGAGATCGCAGGAGTGAAGACACCAACCATCTTTCGCAGCAAGC 539
Db 249 AGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301
QY 540 TCAAGCATCTCAAGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599
Db 302 TCAAGCATCTCAAGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 361
QY 600 CTAACGAGTCTTCAAGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659
Db 362 CTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
QY 660 TCTGTCAAGACATCTCTGAGAGACATACAGAAAGTCTTCTCTCAAGAAAGAAAGCA 719
Db 422 TCTGTCAAGACATCTCTGAGAGACATACAGAAAGTCTTCTCTCAAGAAAGAAAGCA 481
QY 720 GGAGCCGAGTCTATCTTCAACATTTGGCAGAGAGTATTTGCAACATGATTTTCCAGAGT 779
Db 482 GCAACGAGTCTTACCTGCAACCTTGGCAGAGAGTATTTGCAACATGATTTTCCAGAGT 541
QY 780 TTGAACGCTGATGTTTCAACATTTGCAACATTTGCAACATTTGCAACATTTGCAACAT 839
Db 542 GTGAGGAGTGAACATTTGCAACATTTGCAACATTTGCAACATTTGCAACATTTGCAACAT 601
QY 840 GGAATCTTTGGAAG 899
Db 602 GGAATCTTTGGAAG 661
QY 900 GAGTTCAGTGGAGGATTCAGAGATCTCTTGGTGAAGAGGTTTAAATATGTTACG 959
Db 662 GTGCCCGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 721
QY 960 AGGAAGATGAAAGAAATCTTGGGCTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1019
Db 722 AGGAAGATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 781

Qy	1020	TCAGTACCCCTTCTGAACACAGACAGCAGCCATTCGCCAAGAGCAGGAGAAAGGGCGAGGCTTG	1079
Db	782	TCAGCAGCGCTCCTCAAGCAGACAGCAGCCACTGCCAAGAGCGGAGAGCGGCGAGACTGG	841
Qy	1080	AGGACGCTTCATCTATGCTGTGATGAAGAGAGATGATTTCTACATGTTTACTACTCT	1139
Db	842	AAGATGCTTCATCTTATGCTGTGACCAAGGACAGGACTTTCTAAATGTTTACTACTCT	901
Qy	1140	TCCCTAAGAACACACCTCCCTGATTTCTCCCGCATCATAAAGCGAGCTGCTCAGCTAT	1199
Db	902	TCCGGAAGAACACACAGCCCTGCTTCTCCCTGGATCATAAAGCGGCTGCTCGCATAC	961
Qy	1200	TATATGAAGCGAGTGAAGTTCGTTAATGCTCCCTGCTTCCATTAATGATTCGACGG	1259
Db	962	TGTACGAAGACGACGCTGAGGTGCTCCTGATGCCCTGCTTCCGAAGTGAAGTACCG	1021
Qy	1260	AGTTTGTGATCAGCCCTACTTGTGTACTGCTGCTTACCTGCTTACATGAAGACCAAGCCATCC	1319
Db	1022	AGTTTGTGAACCAAGCCCTATTTGCTACTCCGTTTCATGTGAAGACCAAGCCGCTCC	1081
Qy	1320	TGTCGCCACGAAACCCAGTCCCTGCTGCTGATTCACATGCTATTTCTGCAAGACAT	1379
Db	1082	TGTCGCCACGAGCCCGAGTCTGCTGATGATGCCGCTTCTGCTTCTGCAAGACTT	1141
Qy	1380	TTCCATTTCCATTTGATGACAAAGATATGACAATTCGCAATTTGGCAATGGCATCA	1439
Db	1142	TCCCGTTCCATTTTCATGTCGACGAGACCTGCGCATCTGCAAGTGGTAAAGGCATCA	1201
Qy	1440	GAAGCTGTATCAGACAGAGACTTTCAAGGAAGCCCTAATTTGAAGATACTTTGAAA	1499
Db	1202	GAAGCTGTGTGAACAGAGGACTTCCAGGGGAAGCCCACTTTGAAGAGTTCTTTGANA	1261
Qy	1500	TTCTGACTCCAAAATCAACACAGACTTTAGCGGGATCATGACTATCTTGAATATGCAGT	1559
Db	1262	TTCTAATCCCAATCAACAGACATTTAGTGCATCATGACATGTTGAACATGCAGT	1321
Qy	1560	TTGTTGTACGAGTGAAGATGGACAACCTCTGTGAAGAAATCTTCAAGGGTTATGGACC	1619
Db	1322	TTGTATCCCGGTGAGGAGATGGATAACTCGGTGAAGAAATCTCAAGGGTTATGGATC	1381
Qy	1620	TCAAGGCCAATGATCATATGTTGAATCCAGTGAATCTGTTTGGGTGACCCCT	1679
Db	1382	TCAAGGTCAATGATCATATGTTGAATCCAGTGCATCTGTTTGGGTGACCAT	1441
Qy	1680	GTGTGACAGATTAGAAGATTTACAGGACGAGGCTCTACCTCTCAGACATCCCAATTC	1739
Db	1442	GTGTGGACAGCTGGAAGATTCACAGGAGGGGCTCTATCTGTCCGACATCCCAATTC	1501
Qy	1740	ACAATGACCTGAGGATGTGCTTTAATAGGGAAACAGCCCGAGCTCAAGATGGCCTGA	1799
Db	1502	ATACGCCCTGAGGATGTGCTTTGATAGGGAGCAGGACGGCTCAAGATGGCCTCA	1561
Qy	1800	AGAAGAGCTGGGAAGCTGAGGCTACCTTGAAGCAAGCCCAAGCCCTGGAGAGG	1859
Db	1562	AGAAGAGTTGGGAAGCTGAAGCAAGCCCTGGAGCATGCCCAAGCCCTGGAGGAG	1621
Qy	1860	AGAAGAAAAGACAGTACAGCTTCTGTGCTCCATATTTCCCTGTGAGGTTGCTCAGCAGC	1919
Db	1622	AGAAGAGAGACAGTGAATCTGCTGTCTATCTTCCCTCTGAGGTTGCTCAGCAGC	1681
Qy	1920	TGTGGCAAGGCAAGTTGTGCAAGCCCAAGATTTAGTATGTCACCATGCTCTTCTCAG	1979
Db	1682	TGTGGCAAGGCAAAATTTGTCAAGCCCAAGAAATTCAGCGAGGTACCATGCTTTCTCAG	1741
Qy	1980	ACATCGTTGGTTACATGCCATCTGCTCCAGTGCCTCAGCTGCTCAGGCTCATCAGATGC	2039
Db	1742	ATATCGTAGGGTTACATGCTCTATCTGCTCTCAGTGTTCACCTCTGAGGTCATCAGATGC	1801
Qy	2040	TCAATGACATGATACATCGCTTGACACAGCAGTGTGAGAGCTGGATGTCTACAAGTGG	2099
Db	1802	TCAACGCTCTACATCGCTTTGACACAGCAGTGTGAGAGCTGGATGTCTACAAGTGG	1861
Qy	2100	AGACCATTTGCGGATGCCCTATTTGTAGCTGGGGATTACACAAGAGAGTATCATCTCATG	2159
Db	1862	AGACCATCGGGGATGCATATTTGTGTCAGAGTGGATTGCACAGAGAGTGCACACCATG	1921
Qy	2160	CTGTTTCAGATAGCGCTGATGGCCCTGAAGATGATGGAGCTCTCTGATGAAGTATGTCCT	2219
Db	1922	CTGTCCAGATAGCAGCTGATGGCCCTGAAGATGATGGAGCTCTCCAATGAGGTCTATGTC	1981
Qy	2220	CCCATGAGAACCTATCAAGATGGAATTTGAGTGCATCTGGATCATGTTTGTCTGGCG	2279
Db	1982	CCCAAGGAGAACCTATCAAGATGGAATTTGAGTGCATCTGGATCATGTTTGTCTGGAG	2041
Qy	2280	TCGTTGAGTAAATCCCGTACTGCTTTTGGAAAATGTCACCTCTGCTGCTTAACA	2339
Db	2042	TTGTGCGAGTAGAGTCCCGGCTATTTGCTGTTTGGAAAATGTCACCTCTGGCTAACA	2101
Qy	2340	AATTTGAGTCTGAGTGTACCAAGAAAATCAATGTCAGCCCAACAACTTACAGATTAC	2399
Db	2102	AATTTGAATCTGAGTGTCCCTCGGAAAATCAATGTCAGCCCAACCAATACAGGTTAC	2161
Qy	2400	TCAAAGACTGCTGCTGTTTACCTCCATCAAGGGAGGAACTTCCACCAAACT	2459
Db	2162	TCAAAGACTGCTGCTGTTTGTTCACCCGAGATCAAGGGAGGAGCTTCCACCAAACT	2221
Qy	2460	TCCTAGTGAATCCCGGAACTGCGCATTTCTGGATGCTT---ACCACAAAGAACAA	2516
Db	2222	TCCTAGTGAATTCCTGGGATCTGTCACCTTCTGGATGCTTATCACCATCAAGGACCTA	2281
Qy	2517	ACTCAAAACCATGCTTCCAAAAGAAAGATGGAAGATGGAAGTGCCTTATTTAGGCA	2576
Db	2282	ATTCAAAACCATGCTTCCAGGATAAGATGGAAGATGGAAGTGCCTTATTTAGGCA	2341
Qy	2577	AAGCATCAGGATAGATTAGCAACCTATATACCTATTTAT	2616
Db	2342	AAGCTCAGGGTATAGTATGAGCCACATGCTCTATGT	2381
RESULT 10			
RATGCSB			
LOCUS	Rat soluble quanylate cyclase 82kDa subunit mRNA, complete cds.		
DEFINITION	RATGCSB 2668 bp mRNA linear ROD 27-APR-1993		
ACCESSION	M57405 M36075		
VERSION	M57405.1 GI:204277		
KEYWORDS	quanylate cyclase.		
SOURCE	Rat lung, cDNA to mRNA.		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 2668)		
AUTHORS	Nakane, M., Arai, K., Saheki, S., Kuno, T., Buechler, W. and Murad, F.		
TITLE	Molecular cloning and expression of cDNAs coding for soluble quanylate cyclase from rat lung		
JOURNAL	J. Biol. Chem. 265 (28), 16841-16845 (1990)		
MEDLINE	91009100		
PUBMED	1698769		
FEATURES	Location/Qualifiers		
source	1..2668		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
	/tissue_type="lung"		
	/dev_stage="adult"		
	/tissue_lib="lambda-gtl1"		
	457..2529		
	/EC_number="4.6.1.2"		
	/codon_start=1		
	/product="quanylate cyclase"		
	/protein_id="AAA41206.1"		
	/db_xref="GI:204278"		
	/translation="MFCRKFQDKLTGEGPFLSAPQGVPTPEIEVAVGSQCATLPTCOEFAENAGSHSPQRNRYLHTLAESIGKLIFFPERNLALOPTLAKHKIKENRNSSEKEDLERIIAEBAIAGVPVEVLKDSLELFKICYEDHEILGVGGTLKDLFNSFSTLLKQSSHCQAEFRNRGRLEDAISILCLDKQDFLNVYFFPRRTALLPLGI		
CDS			

KAARILLYESHVEVSLMPPCFRSECTEFVNOPYLLLYSVHYKTSKPSLSPCKPOSSIVLI
 PSLFCMTFFMFLMDLAILQLGNGIRRLKMDQGPNGPFEFFELITPKINQTF
 TSGTILKPPFVIRVRNDVYLKQMIYIVESSALFFGSCPTZKINQTF
 TGRGLSLZTHNALRDVYLGEARQADKKRLKAKLTAHQAHEERKTKL
 DLCLSTPSEVAQLQWQGOIVAKENFTWMLFGLVGTACISQSPLOAVITLNAL
 YTRFDQGDGLDYKVEITGDYCVAGGJHRESHTQVIALKAMKMLSENMWSPH
 GEPIKWRIGLHSGSVAGVYGVKMPRYCLFGNNVTLANKEPSCSVPRKIVSVPTVYR
 LKDCPGFVTRSRRELPPNFPSDIPGCHGLDAYQHGQGNPSKFWFOCKDAEDGNANE
 LGKASGVD*

BASE COUNT	684 a	677 c	676 q	631 t
BASE COUNT	684 a	677 c	676 q	631 t

```
Query Match      51.0%; Score 1536.8; DB 10; Length 2668;
Best Local Similarity 82.5%; Pred. NO. 0;
Matches 1796; Conservative 0; Mismatches 371; Indels 10;
```

Qy	443	AGTGGCTTCGTGTTGTGACGTCTCATATAAGAACTACAGCTCATCAGGAGGAGATCGACG	502
Db	380	AGTAGAGCAGATCTTTCATCAGTCCACATCAACACCGGCTTAATAAGGAGAAACCACTGC	439
Qy	503	AGGGTAAGAGACACCAACACCATGTTCTGCACCAAGCTCAAGGATCTCAGATCACAGGA	562
Db	440	---CAGCTCCAGGAACACCATGTTCTGCAGGAAGTTCAAAGATCTCAAGATCACGGGG	495
Qy	563	GAGTGTCTTTCCTTACTGGCACCAGGTCAAGTTCTCTAACGAGTCTTCAGAGGAGGCA	622
Db	496	GAGTGTCTTTCCTTACTGGTCTCTGGTCAAGTCTTCTACGGAGCCATAGAGGAGGTG	555
Qy	623	GCAGGAAGCTCAGAGAGCTCAANGCAACCGTCCCATCTGTCAAGACATTCCTGAGAG	682
Db	556	GCCGGAGTCTCAGAGCTGCCAGGCCATCTGCCACCTGCCAGGAGTTTGCTG---AG	612
Qy	683	AACATACAAGAAAGTCTCTCCTCAAGAAAAACAGTCGGAGCGGAGTCTATCTTCACACT	742
Db	613	AATGCAGAAGGGAGTCAACCCCAAGAAAGACAAGCGCAACAGAGTCTACCTACACACC	672
Qy	743	TTGGCAGAGAGTATTTGCAACATGATTTTCCAGAGATTTGAAACGGCTGAATGTTGACATT	802
Db	673	CTGGCAGAGAGTATTTGGCAAGCTCATTTTCCCAAGAGTTTGAGAGACTGAACCTTGCACCT	732
Qy	803	CAGAGAACATTTGGCAAGCAACAAATAAAGAAAGCAGGAATCTTTGGAAGACAGAAAGAC	862
Db	733	CAGAGAACATTTGCCAAGCACAATAAAGAAACAGGAATCTTCAGAAAAAAGAGAC	792
Qy	863	TTTGAAAAACAATTCAGAGCAAGCAGTTGTCAGCAGAGAGTTCAGTGGAGGTATCAAA	922
Db	793	CTTGAAGAATAATTCAGAGAAGCAATTTGCGCAGGTCTCCAGTGGAGGTGCTCAAA	852
Qy	923	GAATCTCTTGGTGAAGAGTTTAAATATGTTACGAGGAAGATGAAACATCCTTGGG	982
Db	853	GACTCTCTGGGTGAGGAACCTGTTCAGAGATCTGTTATGAGGAAGATGACACATCCTGGGC	912
Qy	983	GTGTTTGGAGCACCCCTTAAAGATTTTTTAAACAGCTTTCAGTACCCCTCTCGAAACAGAGC	1042
Db	913	GTGTTTGGCGCACCCCTGAAGACTTCTCAACAGCTTCAGCAGCTCCTCGAAGCAGAGC	972
Qy	1043	AGCCATTTGCCAAGACAGAGAAAAGGGGACGGCTTGAGGACGCTCCATTCATGGCTG	1102
Db	973	AGCCACTGCCAGGAGCGGAGAGCGGGGGCGACTGGAAGATGCTCCATCTTGTGGCTG	1032
Qy	1103	GATAAGGAGGATGATTTTACATGTTTACTACTTCTTCCTTAAGAGAACCACTCCCTG	1162
Db	1033	GACNAGACCAAGACTTTCTTAATGTTTACTTCTTCCCGAAGAAACCAAGCCCTG	1092
Qy	1163	ATTCTTCCCGCATCATAAAGGCAGCTGCCTCACGTATTATATGAAACGGAGGTGAGAGT	1222
Db	1093	CTTCTCCCGGTATCATTAAGCGGCTGCTCGCATCTGTACGAAAGCCACGTGGAGGTG	1152
Qy	1223	TCGTTATAGCTCCCTGCTTCCATATGATTCAGCGAGTTTGTGANTCAGCCCTACTTG	1282
Db	1153	TCCCTGATGCTCCCTGCTTCGGAAGTGAGTGTACCGAATTCGTGTGAACCAAGCCCTATTG	1212
Qy	1283	TTGTACTTCGCTTCACATGAAAGACCAAGCCATCCTCTGTCGCCCAAGAAACCCAGTCC	1342

[illegible]

Db 2293 CGGAAATCAATGTACGCCCCACGACATACAGGTACTCAAGACTGTCTGGCTTTGTG 2352

Qy 2423 TTTACCCCTCGATCAAGGGAGGAATTCACCAAACTTCCTAGTGAATCCCGGAATC 2482

Db 2353 TTTACCCCGAGATCGAGGGAGGACTTCCACCAAACTTCCTAGTGACATCTCTGGATC 2412

Qy 2483 TGCCATTTTCTGGATGCTTACCAGAA--CAAGGAAACAACTCAAAACCATCTTCCAAAG 2539

Db 2413 TGTACATTTCTGGATGCTTATCAACATCAAGGACCTAATCTAAACCATGTTTCCAGCAG 2472

Qy 2540 AAGATGTGAAGATGCAATGCCATTTTATAGGCAAGCATCAGCAATAGATTAGCAA 2599

Db 2473 AAGATGCAAGATGGAACGCTAATCTTATAGGCAAGCATCAGGGGTAGATTAGTGA 2532

Qy 2600 CCTATATACCTATTAT 2616

Db 2533 GCCGCAATGTTCTGTGT 2549

RESULT 11

AF233746S1 1571 bp DNA linear PRI 16-APR-2000

LOCUS Homo sapiens soluble guanylate cyclase large subunit gene, exon 3.

DEFINITION AF233746

ACCESSION AF233746.1 GI:7576897

KEYWORDS 1 of 6

SEGMENT Homo sapiens.

SOURCE Homo sapiens.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE Zhou,Y., Zheng,J.B., Gu,X., Li,W. and Saunders,G.F.

AUTHORS A novel Pax-6 binding site in rodent B1 repetitive elements: coevolution between developmental regulation and repeated elements?

TITLE Gene 245 (2), 319-328 (2000)

JOURNAL MEDLINE 20183694

PUBMED 10717483

REFERENCE 2 (bases 1 to 1571)

AUTHORS Zhou,Y.H.

TITLE Direct Submission

JOURNAL Submitted (14-PEB-2000) Biochemistry and Molecular Biology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Bldg., Houston, TX 77030, USA

FEATURES

source 1. .1571

exon /organism="Homo sapiens"

378..1087

/db_xref="taxon:9606"

/number=3

BASE COUNT 517 a 250 c 310 g 494 t

ORIGIN

Query Match 23.88; Score 716.2; DB 9; Length 1571;

Best Local Similarity 99.68; Pred. No. 8.2e-184;

Matches 718; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 897 CAGAGTCCAGTGGAGGTTATCAAGAAATCTCTTGGTGAAGAGGTTTTTAAATATGTT 956

Db 375 CAGAGTCCAGTGGAGGTTATCAAGAAATCTCTTGGTGAAGAGGTTTTTAAATATGTT 434

Qy 957 ACAGGAGATGAACATCCTTGGGTGTTGGAGCACCCTTAAGATTTTTTAAACA 1016

Db 435 ACAGGAGATGAACATCCTTGGGTGTTGGAGCACCCTTAAGATTTTTTAAACA 494

Qy 1017 GCTTCAGTCCCTTCTGAACAGAGCAGCATTCGCAAGAGCAGGAAGAGGCGAGGC 1076

Db 495 GCTTCAGTCCCTTCTGAACAGAGCAGCATTCGCAAGAGCAGGAAGAGGCGAGGC 554

Qy 1077 TTGAGGACGCTCCATCTATGCTGTGAAGAGGATGATTTTCTACATGTTTACTACT 1136

Db 555 TTGAGGACGCTCCATCTATGCTGTGAAGAGGATGATTTTCTACATGTTTACTACT 614

Qy 1137 TCTTCCTAGAGAACACCTCCTCCTGATCTTCCCGGCATCATAAAGGACGCTGCTCACG 1196

Db 615 TCTTCCTAGAGAACACCTCCTCCTGATCTTCCCGGCATCATAAAGGACGCTGCTCACG 674

Qy 1197 TATTATATGAACGGAAGTGAAGTGTCTGTTAATGCTCCTGCTTCCATTAATGATGCA 1256

Db 675 TATTATATGAACGGAAGTGAAGTGTCTGTTAATGCTCCTGCTTCCATTAATGATGCA 734

Qy 1257 GCAGAGTTGTGAATCAGCCCTACTTGTGTACCTCGTTACATGAAAGCACCAGCCAT 1316

Db 735 GCAGAGTTGTGAATCAGCCCTACTTGTGTACCTCGTTACATGAAAGCACCAGCCAT 794

Qy 1317 CCTGTGCCCCCAGCAAAACCCAGTCTCGCTGGTGTATCCACATCGTATTTCTGCAAGA 1376

Db 795 CCTGTGCCCCCAGCAAAACCCAGTCTCGCTGGTGTATCCACATCGTATTTCTGCAAGA 854

Qy 1377 CATTTCCATTCATTTTCATGTTTTCACAAAGATATGACAAATCTTGCAATTTGGCAATGGCA 1436

Db 855 CATTTCCATTCATTTTCATGTTTTCACAAAGATATGACAAATCTTGCAATTTGGCAATGGCA 914

Qy 1437 TCAGAGGCTGTGAACAGGAGAGCTTTCAAGGAAAGCCCTAATTTTGAAGAATACVTTG 1496

Db 915 TCAGAGGCTGTGAACAGGAGAGCTTTCAAGGAAAGCCCTAATTTTGAAGAATACVTTG 974

Qy 1497 AAATTCTGACTCCAAAATCAACACAGACGTTTACGGGATCATGACTATCTTGAATATGC 1556

Db 975 AAATTCTGACTCCAAAATCAACACAGACGTTTACGGGATCATGACTATCTTGAATATGC 1034

Qy 1557 AGTTTGTGTACAGTGAAGAGATGGGACAACTCTCTGAAGAAATCTTCAAGGGTTATGC 1616

Db 1035 AGTTTGTGTACAGTGAAGAGATGGGACAACTCTCTGAAGAAATCTTCAAGGGTTATGC 1094

Qy 1617 A 1617

Db 1095 A 1095

RESULT 12

AC021433/c

LOCUS

DEFINITION Homo sapiens chromosome 4 clone RP11-688K17 map 4, WORKING DRAFT

SEQUENCE, 12 unordered pieces.

AC021433

AC021433.3 GI:7417774

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

1 (bases 1 to 177556)

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

2 (bases 1 to 177556)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F., Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N., Pisanic,C., Pollara,V., Raymond,B., Riley,R., Rothman,D., Roy,A., Santos,R., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo.A., Wu.X., Wyman,D., Ye.W.J., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

AB062171 3022 bp mRNA linear VRT 19-JUL-2001
 Locus Takifugu rubripes FRGCS-alpha mRNA for soluble guanylyl cyclase
 DEFINITION alphan subunit, complete cds.
 ACCESSION AB062171
 VERSION AB062171.1 GI:14495183
 KEYWORDS Takifugu rubripes cDNA to mRNA.
 ORGANISM Takifugu rubripes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Takifugu.
 1
 Morinaga, C., Yamamoto, T., Moriya, Y. and Suzuki, N.
 Identification of tandem organization of soluble guanylyl cyclase
 alphan and beta subunit genes in the Japanese pufferfish (Fugu
 rubripes) genome
 Unpublished
 2 (bases 1 to 3022)
 Morinaga, C., Yamamoto, T., Moriya, Y. and Suzuki, N.
 Direct Submission
 Submitted (23-MAY-2001) Norio Suzuki, Hokkaido University, Division
 of Biological Sciences; Kita 10, Nishi 8, Kita-ku, Sapporo,
 Hokkaido 060-0810, Japan (E-mail: norio-suzuki.hokudai.ac.jp,
 Tel: 81-11-706-4908, Fax: 81-11-746-1512)
 Location/Qualifiers
 1. 3022
 /organism="Takifugu rubripes"
 /db_xref="taxon:31033"
 /note="synonym: Fugu rubripes"
 1. 3022
 /gene="FRGCS-alpha"
 226..2253
 /gene="FRGCS-alpha"
 /codon_start=1
 /product="soluble guanylyl cyclase alphan subunit"
 /protein_id="BAB60907.1"
 /db_xref="GI:14495184"
 /translation="MFCTKLKELKISGCPFTSSAKTNQLEHSQSLKDVHLPSTKSD
 VRRNTGNAPHOXSRAKYNLHTGISIRKLACPEFERLHTALRMRLSDPSREPAS
 PFCVCTEYQSCDDEHFVWMSFSSKAIOMDTLVALGEELEFNWYEDRHILRV
 VGGALHEDANSFVLLKQSLTHYQREDVCNERSVLCIDKDLGLLVYFNPHPTTE
 LFFGVVKAARALVDVDMDDPATKDSILQSPSLLYTVVYKDKSLSPSL
 RATSAGTFLSFTIPFPHLIDLDLVQVGHGLRKLRTKDLRSPATFQSHFSI
 LSPILRTFCQILFMLNQTIRIKRGVSTADNTLMDLKGOMIYVPSDAILFTGSPC
 VKLEELTGRGLYLSLIP INNALRDVVLVGEQAKADGLKRLGKAKALEHAHOALE
 EEKKTVDLLFTIPGTVAOQLHOGQTVQAKPERVTLPSDIDVFTAVGSHCTPMQV
 ITMLNELYTFDHHGGLDVKVETIGDAICVAGLHKESETHAVQVLMALMKWELLS
 NGVMTAPGEPIQMRIGLHTGSVLGVVGVKMPRYCLFGNNVTLANRKFESCSQPKINI
 SPVTHRLLEDHPEVFIPRRRQELPANFPEDIPGVCFLEASPRPSEVTLK"
 BASE COUNT 769 a 766 c 726 g 761 t
 ORIGIN
 Query Match 21.9%; Score 660.2; DB 5; Length 3022;
 Best Local Similarity 64.7%; Pred. No. 1.8e+168;
 Matches 1034; Conservative 0; Mismatches 553; Indels 12; Gaps 3;
 QY 909 TGGAGGTTATCAAGAATCTCTGGTGAAGAGGTTTTTAAATATGTTACGAGGAAGATG 968
 DB 629 TGGTACTCTGAGTAGCGCTTGGCAGGAGCTTTTCAACATGTGCTATGAAGAGACA 688
 QY 969 AAAACATCTTGGGTGGTGGAGCCACCTTAAAGATTTTTTAAACAGCTTCAGTACCC 1028
 DB 689 GGCACATTTTGGGTGGTGGGAGCCCTGACGACGCTCTCTCAACAGCTTCAATGCT 748
 QY 1029 TTCTGAACAGCAGCAGCATTTGCCAAGAGCAGGAAAAGGGCAGGCTTCAGCAGCCT 1088
 DB 749 TGGTGAACAGCAGCAGCAGCTGCTATACCAGGACAGAGAGGATTTGTATAACCAACTT 808
 QY 1089 CCATTCATCTGCTGGAAGAGGATGATTTTCTACATGTTTACTACTTCTTCCCTAAGA 1148
 DB 809 CGGTGCTGCTTAGACAGGATTTGGTCTGCTTACTGCTATTTCTTCAATCCCCACC 868

1149 GAACCACTCCCTGATTCTTCCCGGCATCATAAAGCAGCTGCTCAGTATATATGAAA 1208
 DB 869 CGACCACTAGCTCTTTTCCCGGAGCTCATAAAGCTGCTGCCGCTGCTCTATCACA 928
 QY 1209 CGGAAGTGGAGTGTCTGTTAATGCTCCCTGCTTCCATATATGATTGAGCAGGTTGTGA 1268
 DB 929 CCACGTGGATGTTCTGTATGACCTCTCTGCACTAAAGATAGCATCTCGACGACGAC 988
 QY 1269 ATCAGCCCTACTTCTGTACTCCCTTCCATGAAAGCAGCAAG--CCATCCCTGTCCTC 1325
 DB 989 CGCAGCCAGTCTTCTGTACACAGTTGTAGTTAAAGATGCTTAAAGTCTTGAGCCCCAGTC 1048
 QY 1326 CCAGCAAAACCCAGTCTCTGCTGCTGATTTCCACATGCTATTCTTCCAAAGACATTTCCAT 1385
 DB 1049 CACTACAGCCACCTCAGCGGGAGCTTCTTACTCTCTGTTCTCCACCATCTTCCCTT 1108
 QY 1386 TCCATTTCAATGTTTGACAAAGATGACAAATCTTGCATTTGGCAATGGCAATGAGCAGG 1445
 DB 1109 TCCATCTGATCTCGACGAGCTTGTCTGTGCAAGTAGGACACGGGCTGAGGAAGA 1168
 QY 1446 TGTGAACAGGAGAGA--CTTTCAAGGAAGCCTAATTTTGAAGAATACTTTCAAAATTC 1502
 DB 1169 GACTCACAGGAAAGATGGAGCTGGCAGCAGCAGCCACTTCCAAAGACACTTCTCCATTC 1228
 QY 1503 TGACTCCAAAATCAACCCAGACGCTTTAGCGGATCATGACTATCTTGAATATGCAAGTTG 1562
 DB 1229 TTTCTCCCTAATCAGATGATCTTCCAAGGTATTTCAACCATCTGACACTCAGTTCA 1288
 QY 1563 TTGTACAGTGGAGATGGGACAACTCTCTGTGAAGAAATCTTCAAGGGTTATGAGACCTCA 1622
 DB 1289 CCATTCGGATCAAGCGTGGAGTCTCCACCGCAGATAACAC-----GCTATGAGCCTTA 1342
 QY 1623 AAGCCAAATGATCTACATTTTGAATCCAGTGAATCTTGTGTTTGGGGTCAACCTGTG 1682
 DB 1343 AAGTCAAGATGATTTATGTCAGAGTCCGAGCCATTTGTTTGGGCTCGCCGCTGTG 1402
 QY 1683 TGGACAGATTAGAAAGATTTTACAGGACGAGGCTTCTACCTCTCAGACATCCCAATTCACA 1742
 DB 1403 TAGCAAGCTGGAGGAGCTGACGGCTGCTGSCCTGTACCTGTGACATCCCATTCATA 1462
 QY 1743 ATGCACGTGAGGATGTGGTCTTATAGGGGAACAAGCCGAGCTCAAGATGGCCCTGAAGA 1802
 DB 1463 ACGGCTGCGCGAGTGTCTGCTGTTGGGCAACAGGCAAGCCAGGAGGCTGGAAGA 1522
 QY 1803 AGAGCTGGGGAAGCTCAAGGCTACCTTGAGCAAGCCACCAAGCCCTGGAGGAGGAGA 1862
 DB 1523 AGCGCTGGGGAAGCCCAAGCAGCCCTTGGACATGCCCATCAAGCACTGGAGGAGGAGA 1582
 QY 1863 AGAAAAAGACAGTAGACCTTCTGTGCTCCATATTTTCCCTGTGAGGTTGCTCAGCAGCTGT 1922
 DB 1583 AGAAGAAGCCGTAGAGCTCTCTCTTCAACATCTTTTCCCGGCACCGTTGCGCAGCAGCTGT 1642
 QY 1923 GGCAGGCAAGTGTGCAAGCCCAAGATGCTCAATGTCACCATCTCTTCTCAGACA 1982
 DB 1643 GGCAGGCCCAGACGGTCCAGGCAAGAGTTTGAAGGGTCAACAATGCTCTTCTCTGACA 1702
 QY 1983 TCGTGGGTTCACTGCCATCTGCTCCAGTGTCTCACCGCTCGAGGCTCATCACCATGTCTCA 2042
 DB 1703 TCGTGGGTTTTCACGCGGTTTGTCTCTCACTGCACCCCGATGCAAGTATCACCATGTGA 1762
 QY 2043 ATGCACGTACACTCGCTTGCACAGCAGTGTGGAGAGCTGGATGCTTACAAGTGGAGA 2102
 DB 1763 ACGAGCTGACACCAAGTTTGACCAACCTCTGGAGAGCTGGAGCTGTATAAGTGGAGA 1822
 QY 2103 CCATTCGGATGCTTATTTGTAGCTGGGATTTACAAAGAGAGTGTACTCTATCTGCTG 2162
 DB 1823 CCATTCGGTGTGCTTATTTGCTAGCCCGGCTTTACACAGAGGAGGAGACTCAGCGCTG 1882
 QY 2163 TTCAGATAGGCTGATGCCCTTGAAGATGATGGAGCTCTCTGTATGAAGTATGTCTCCCC 2222
 DB 1883 TCCAAGTAGCAGTCTATGGCTTAAAGATGATGGAGCTTTTCGAACGGGTGTTATGACTCCAG 1942

Db 1461 GGATGGCTTGAAGAAGCGGCTGGGAAAGCCAGGCGCCTTGAAACGCTCACCAGC 1520
Qy 1849 CCTGGAGGAGAGAGAAAGACAGTAGACCTTCTGTGTCTCCATATTTCCCTGTGAGGT 1908
Db 1521 GCTGGAGGAGAGAGAGAGACAGTAGATCTCTCTTTCCATTTTCCAGGCACCGT 1580
Qy 1909 TGCTCAGCAGCTGTGGCAAGGCAAGTGTGCAAGCCAAAGATTCAGTAATGTCCACAT 1968
Db 1581 GGCCACAGCAGCTGTGGCAGGGGAACTGTTCAAGCAAGAAAGTTAAACCAAGTTACGAT 1640
Qy 1969 GCCTTCTCAGACATCGTTGGGTTCACTGCCATCTGCTCCAGTGTCCACCGTGCAGGT 2028
Db 1641 GCTCTTCTCGATATTGGGCTTCACAGGGTTTGTCCCTCTGCACACCAATGCAAGT 1700
Qy 2029 CATCACCATGCTCAATCAGCTGTACACTGCTCGACCAGCAGTGTGGAGAGCTGGATGT 2088
Db 1701 GATCACCATGCTCAATGAGCTGTACACTAAATTTGACTATCAATGCCGAGAGCTTGATGT 1760
Qy 2089 CTACAAGGTGGAGACCATTGGCGATGCTATTGTGTAGCTGGGGGATTACACAAAGAGAG 2148
Db 1761 CTATAAGGTAGAGACCATTAGTGATGCTTATTGTGTAGCCGTTGCACAGGGAGAG 1820
Qy 2149 TGATACACTACGCTGTTTCAAGTAGCGCTGATGGCCCTCAAGATGATGGAGCTCTCTGATGA 2208
Db 1821 TGAACCCATGCACTGCAATTCGATTCATTCATGGCAATGGAAGATGATTGAGCTTTCAGATGA 1880
Qy 2209 AGTTATGCTCCCATGGAACCTATCAAGATCGCAATTTGGACTGCACCTCTGGATCAGT 2268
Db 1881 AGTTTTCACACCAACCGAGAGCCAAATACAGATGCGTATCGGTCTCCACAGTGGTCACT 1940
Qy 2269 TTTTGGCTGGCTGCTGGAGTTAAATGCCCGTTACTGTCTTTTGGAAACAAATGTAC 2328
Db 1941 GTTGGCTGGTGTGTTGGGTTAGGATGCCAAGGTACTGCTTGTGTTGAAACAATGTTAC 2000
Qy 2329 TCTGGCTAAACAATTTGAGTCTCCTGAGTGTACACGAAATCAATGTGAGCCCAACAAC 2388
Db 2001 ATTAGCCCAACAAGTTTCAGTCATGTCAGCCAGCCGAGAAAGTCAACATCAGCCCAACAAC 2060
Qy 2389 TTACAGATTACTCAAGAGCTGCTCTGTTTTCGTTTACCCCTCGATCAAGGGAGGAAC 2448
Db 2061 GCACAGATTGGTCAAGGTCGCCAGAGTTTGTGTTTATTTCCAGGAGCAGACAAGACCT 2120
Qy 2449 TCCACCAACTTCCTTAGTGAATCCCGGAATCTGCCATTTTCTGGATGCTTACCAACA 2508
Db 2121 TCCGGCCAACCTTCCCAAGACATCCCTCGTGTGTTTACTTTTGGAGTCAGCCTTGAA 2180
Qy 2509 AGGAACAAA 2517
Db 2181 AATGTCAAA 2189

Search completed: July 1, 2003, 16:02:28
Job time : 5176.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 10:46:42 ; Search time 4186.4 Seconds
(without alignments)
16983.119 Million cell updates/sec

Title: US-09-762-767A-3
Perfect score: 2443
Sequence: 1 cccccccgcgcgtgcgc.....aaatcatttttaaggagt 2443

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rnd.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em_sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2443	100.0	2443	6	AX024601	AX024601 Sequence
2	2443	100.0	2443	9	HSGCSAB	X65533 H.sapiens s
3	1842	75.4	1951	9	AF020340	AF020340 Homo sapi
4	1774	72.6	3137	4	BTGUCYC	Y00770 Bovine mRNA
5	1711.4	70.1	3235	10	AF297083	AF297083 Mus muscu
6	1691.8	69.3	3047	10	RATGCS	M2562 Rat solubile
7	1652.4	67.6	2303	10	AF020339	AF020339 Mus muscu
8	1151.6	47.1	2997	5	AB000850	AB000850 Oryzias l
9	1087.8	44.5	2769	5	AB062172	AB062172 Takifugu
10	582.2	23.8	2593	3	AB062386	AB062386 Hemicentr
11	559.6	22.9	3691	3	AF062751	AF062751 Manduca s
12	521	21.3	70709	9	AC114761	AC114761 Homo sapi
13	521	21.3	177556	2	AC021433	AC021433 Homo sapi
14	452.2	18.5	143255	2	AC015996	AC015996 Homo sapi
15	371.4	15.2	2840	3	DMU27123	U27123 Drosophila
16	270.6	11.1	325	4	AF486295	AF486295 Ovis arie
17	267	10.9	5575	3	AB062387S4	AB062387 Hemicentr
18	218	8.9	299480	2	AC121563	AC121563 Mus muscu
19	213.2	8.7	218756	2	AC128391	AC128391 Rattus no
20	212.4	8.7	8133	3	AF017062	AF017062 Anopheles
21	203.6	8.3	218756	2	AC128391	AC128391 Rattus no
22	203.4	8.3	2657	10	AF109963	AF109963 Rattus no
23	200.2	8.2	2954	6	AX281673	AX281673 Sequence
24	200.2	8.2	2954	9	HSGCA2	X63282 H.sapiens m
25	192.6	7.9	2258	9	HSU58855	U58855 Homo sapien
26	192.6	7.9	2627	9	AK096588	AK096588 Homo sapi
27	192.6	7.9	2657	9	BC028384	BC028384 Homo sapi
28	192.6	7.9	3015	6	AX024599	AX024599 Sequence
29	192.6	7.9	3015	9	HSRNASGC	Y15723 Homo sapien
30	187.2	7.7	2445	10	AF297082	AF297082 Mus muscu
31	186.2	7.6	3004	9	HSGCSAA	X66534 H.sapiens s
32	178.2	7.3	2668	10	RATGCSB	M57405 Rat solubile
33	178.2	7.3	4775	10	RNU60835	U60835 Rattus norv
34	177	7.2	3261	4	BTGUCV73	X54014 Bovine mRNA
35	174.2	7.1	3022	5	AB062171	AB062171 Takifugu
36	167.4	6.9	9762	5	AB062170	AB062170 Takifugu
37	167.4	6.9	17366	5	AB022281	AB022281 Oryzias l
38	160.6	6.6	3047	9	HSRNASCA2	Z50053 H.sapiens m
39	150.4	6.2	2570	5	AB000849	AB000849 Oryzias l
40	149.4	6.1	62894	2	AC113467	AC113467 Mus muscu
41	147.8	6.0	661	10	AB079780	AB079780 Rattus no
42	147.8	6.0	206456	2	AC112402	AC112402 Rattus no
43	147.2	6.0	2335	6	AX401893	AX401893 Sequence
44	147.2	6.0	2335	10	RATGCSB2	M57507 Rattus norv
45	147.2	6.0	2452	10	AY004153	AY004153 Rattus no

ALIGNMENTS

RESULT 1
AX024601
LOCUS AX024601 Sequence 3 from Patent DE19837015.
DEFINITION AX024601
ACCESSION AX024601
VERSION AX024601.1 GI:10184740
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2443)
AUTHORS Poller.W., Schmidt,H. and Zabel,U.
TITLE New human soluble guanylate cyclase alpha/beta1 and the nucleic acid encoding the subunits, useful for producing diagnostic

linear PAT 15-SEP-2000


```
QY 1981 CTAGGTTCCAGTTTCTCCTAACACGTCGCCAAGCCAGGACGAGTCTTCCCTATGGATA 2040
Db 1981 CTAGGTTCCAGTTTCTCCTAACACGTCGCCAAGCCAGGACGAGTCTTCCCTATGGATA 2040
QY 2041 CAGATTCTTTCTTTCTCCTTGTCCTTACCCCAAGACTTCTCTAGATATATCTCTCACT 2100
Db 2041 CAGATTCTTTCTTTCTCCTTGTCCTTACCCCAAGACTTCTCTAGATATATCTCTCACT 2100
QY 2101 ATCCGTTATTCACACTTACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
Db 2101 ATCCGTTATTCACACTTACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
QY 2161 AAGTTTGGCTTTTGTATGGGATGATGAGCTTCTATGCTTCTTAAATCTACTACAGCA 2220
Db 2161 AAGTTTGGCTTTTGTATGGGATGATGAGCTTCTATGCTTCTTAAATCTACTACAGCA 2220
QY 2221 TTACTTAACATGGTGTATGTCGAAGTAGTAGGACCCCAATAATATTTTGAATTTAGTT 2280
Db 2221 TTACTTAACATGGTGTATGTCGAAGTAGTAGGACCCCAATAATATTTTGAATTTAGTT 2280
QY 2281 AAATGAACACTGAACAGTGTGGCCATGTGTATATTTATATCATGTTTACCAAAATCTGTT 2340
Db 2281 AAATGAACACTGAACAGTGTGGCCATGTGTATATTTATATCATGTTTACCAAAATCTGTT 2340
QY 2341 TAGTGTCCACATATATATATATATATATTTTAAATGACTATATATGTAATAGTTTATAT 2400
Db 2341 TAGTGTCCACATATATATATATATATTTTAAATGACTATATATGTAATAGTTTATAT 2400
QY 2401 CATGTTGGTGTATATATATATAGAAATCATTTTCTTAAAGGAGT 2443
Db 2401 CATGTTGGTGTATATATATAGAAATCATTTTCTTAAAGGAGT 2443

RESULT 2
HSGCSAB
LOCUS 2443 bp mRNA linear PRI 19-NOV-1992
DEFINITION H.sapiens soluble guanylate cyclase small subunit mRNA.
ACCESSION X65333
VERSION X65333.1 GI:31685
KEYWORDS cytoplasmic protein; GTP pyrophosphate-lyase; Guanylate cyclase.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Guilla, G., Scholl, U., Bulle, F. and Guellaen, G.
TITLE Molecular cloning of the cDNAs coding for the two subunits of
JOURNAL soluble guanylyl cyclase from human brain
MEDLINE FEBS Lett. 304 (1), 83-88 (1992)
PUBMED 92316204
1352257
REFERENCE 2 (bases 1 to 2443)
AUTHORS Guellaen, G.
TITLE Direct Submision
JOURNAL Submitted (10-JUL-1992) Georges Guellaen, Unite INSERM 99, Hopital
Henri Mondor, 94010 Creteil, FRANCE
FEATURES
location/Qualifiers
1..2443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4q31.3-q33"
/tissue_type="brain, frontal lobe"
/clone_lib="lambda GT10"
89..1948
/EC_number="4.6.1.2"
/notes="soluble"
/codon_start=1
/product="guanylate cyclase"
/protein_id="CAA47144.1"
/db_xref="GI:31686"
/db_xref="SWISS-PROT:Q02153"
/translation="MYGFVNHALLVIRNYPVEWEDIKKEAQIDBEGQFLVRIYD"

CDS
```

```
DSKYDLVAASKVLNLNAGBILQFMKFFVFCQESGYDILRLVLSNNREFQLNLD
ALHDHATIPGMRAFSRCDTAEKGLGLIHYSREGQDIIVIGIKTVAQOIHGT
EIDMKVIOQRNECDHTQFLIEEKSEDEEDLDRFENGQIESRPTFTCKAFKFP
PHIFDRDLVVTQCGNATYLSVLPOLPGNCSLLSVSLVRPHIDISPHGILSHLNTVF
VLRKGLDLVEKLECEDELGTETISCLRLKAGOMYLPEADSLIFLPSVSNLDDLT
RRGLYLSIDIPLDHATRDLYLLGEOPREYKLTLELEILDRQLTLRALEDEKKTDT
LLSYLSPSVANELRHRKRPVPAKRYVDNVITLFGIVGFNAFCSHASGEAGANKIVNLL
NDLYRFDTLDSRKNPEVYKVEIVGDMVTGSLVSGPEPFIHARSICHLAIDMEIAG
QYVDGESVQITIGITHTGEVTVIGQRPYICLFGTNTLSTETTGEGKINVSE
YTYRCLAMSPENDPQFHLHGRGPMKKEPMQVWFLSRKNTGTETTKQDD*
BASE COUNT 709 a 513 c 533 g 688 t
ORIGIN
Query Match 100.0%; Score 2443; DB 9; Length 2443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCCCCCCCCCGCGTCCGCGCTCGCGTGGGTGCGCTTCGGCGGTACCTTCGCGTGGGGC 60
Db 1 CCCCCCCCCCGCGTCCGCGCTCGCGTGGGTGCGCTTCGGCGGTACCTTCGCGTGGGGC 60
QY 61 TGCCTCCCGCGCTCCGCGTGCAGACACCATGTACGGATTTGTGAATCAGCGCCTCGAGTT 120
Db 61 TGCCTCCCGCGCTCCGCGTGCAGACACCATGTACGGATTTGTGAATCAGCGCCTCGAGTT 120
QY 121 GCTGTGATCCGCAATTAACGCCCGAGGTGTGGAGACATCAAAAAGAGGACACGTT 180
Db 121 GCTGTGATCCGCAATTAACGCCCGAGGTGTGGAGACATCAAAAAGAGGACACGTT 180
QY 181 AGATGAAGAAGACAGTTTCTTGTCAGATAATATATATGATGATGATGATGATGATGAT 240
Db 181 AGATGAAGAAGACAGTTTCTTGTCAGATAATATATATGATGATGATGATGATGATGAT 240
QY 241 GGTGCTGCTCAAGCAAGTCTCAATCTCAATCTGAGAGAAATCTCCAAATCTTTGG 300
Db 241 GGTGCTGCTCAAGCAAGTCTCAATCTCAATCTGAGAGAAATCTCCAAATCTTTGG 300
QY 301 GAAGATGTTTTTCTGTTTTGCCAAGAATCTGGTTATGATACAATCTTTCGCTGCTCTGG 360
Db 301 GAAGATGTTTTTCTGTTTTGCCAAGAATCTGGTTATGATACAATCTTTCGCTGCTCTGG 360
QY 361 CTCATATCTCAGAGATTTTCTACAGAACCTTGATGCTCTGCAGACACCTTGCTTACCAT 420
Db 361 CTCATATCTCAGAGATTTTCTACAGAACCTTGATGCTCTGCAGACACCTTGCTTACCAT 420
QY 421 CTACCCAGGAATGCGTGCACTCTTCTTTAGTGTGCTGATGAGAGAAAGGGCAAGGACT 480
Db 421 CTACCCAGGAATGCGTGCACTCTTCTTTAGTGTGCTGATGAGAGAAAGGGCAAGGACT 480
QY 481 CATTTTGCACCTACTACTCAGAGAGAGAGGACTTCCAGGATATTTGTCATTTGGATCATCAA 540
Db 481 CATTTTGCACCTACTACTCAGAGAGAGAGGACTTCCAGGATATTTGTCATTTGGATCATCAA 540
QY 541 AACAGTGGCACAACAATCCATGGACCTGAATAGACATGAGGTTATTCAGCAAGAA 600
Db 541 AACAGTGGCACAACAATCCATGGACCTGAATAGACATGAGGTTATTCAGCAAGAA 600
QY 601 TCAAGAATGTGATCATACTCAATTTTAAATTAAGAAAAAGAGTCAAAAGAGAGAGATTT 660
Db 601 TCAAGAATGTGATCATACTCAATTTTAAATTAAGAAAAAGAGTCAAAAGAGAGAGATTT 660
QY 661 TTATGAAGATCTTGACAGATTTGAAGAAAAATGGTACCAGGAATCAGCATCAGCCCAT 720
Db 661 TTATGAAGATCTTGACAGATTTGAAGAAAAATGGTACCAGGAATCAGCATCAGCCCAT 720
QY 721 TACATCTGCAAGCTTTTCTTTCATATATATTTTACCGGGACCTAGTGGTCACTCA 780
Db 721 TACATCTGCAAGCTTTTCTTTCATATATATTTTACCGGGACCTAGTGGTCACTCA 780
QY 781 GTGTGGCAATGCTATATACAGAGTTCTCCCCAGCTCCAGCTCGGAAATTTGAGGCTTCT 840
Db 781 GTGTGGCAATGCTATATACAGAGTTCTCCCCAGCTCCAGCTCGGAAATTTGAGGCTTCT 840
```


QY 1838 TTGAGCAGAGAGCCAGTGTCCATGAAGGCAAAAGAACCAATGCAAGTTTGTTT 1897
 Db 1702 TTGAGCAGAGAGCCAGTGTCCATGAAGGCAAAAGAACCAATGCAAGTTTGTTT 1761
 QY 1898 CTATCCAGAAAAATACAGCAAGAGGAAACAAAGCAGGATGATGACTGATCTGGAT 1957
 Db 1762 CTATCCAGAAAAATACAGCAAGAGGAAACAAAGCAGGATGATGACTGATCTGGAT 1821
 QY 1958 TATGGGTGAAGAGGAGTACAGACTAGTTTCCAGTTTCTCCTAACACGCTGCCAAGCCCA 2017
 Db 1822 TATGGGTGAAGAGGAGTACAGACTAGTTTCCAGTTTCTCCTAACACGCTGCCAAGCCCA 1881
 QY 2018 GGAGCAGTCTTCCCTATGATACAGATTTCTTTGCTCTGCTATACCCAGACT 2077
 Db 1882 GGAGCAGTCTTCCCTATGATACAGATTTCTTTTGTCTCTGCTATACCCAGACT 1941
 QY 2078 TTCTTCTAGA 2087
 Db 1942 TTCTTCTAGA 1951

BTGUCYC 3137 bp mRNA linear MAM 14-MAY-1991
 LOCUS Bovine mRNA for soluble guanylate cyclase 70 kDa subunit.
 DEFINITION Y00770
 VERSION Y00770.1 GI:407
 KEYWORDS Guanylate cyclase.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 3137)
 Authors Gauseprkl, H., Herz, J., and Koesling, D.
 TITLE Direct Submission
 JOURNAL Submitted (02-DEC-1988) H. Gauseprkl, J. Herz and D. Koesling,
 INSTITUT FUER PHARMAKOLOGIE, FREIE UNIVERSITAET BERLIN, THIELALLEE
 69173, 1000 BERLIN 33
 2 (bases 1 to 3137)
 Authors Koesling, D., Herz, J., Gauseprkl, H., Niroomand, F., Hinsch, K. D.,
 Mulsch, A., Bohme, E., Schultz, G. and Frank, R.
 TITLE The primary structure of the 70 kDa subunit of bovine soluble
 guanylate cyclase
 JOURNAL FEBS Lett. 239 (1), 29-34 (1988)
 MEDLINE 89031214
 PUBMED 2903071

FEATURES
 source
 1..3137
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /tissue_type="lung"
 1..3137
 /note="guanylate cyclase 70 kDa subunit"
 /evidence=experimental
 87..1946
 /EC_number="4.6.1.2"
 /codon_start=1
 /product="guanylate cyclase 70 kDa subunit"
 /protein_id="CA468739.1"
 /db_xref="GI:408"
 /db_xref="SWISS-PROT:P16068"
 /translation="MYGFVNHALLELVIRNYPVEWEDIKKAQLDEEQFLVRIYD
 DSKTYDLVAASVNLNAGEILQMFGRMFVFCQESGYDTILRLGNSVNFPLQND
 ALDHLYATVPGMRAPSFRTDADKGLILHYSEREGLODIVIGIITVAQOIHT
 EIDMKVIQORNECDHTQPLIEKESKEDEVEDLDFRENGTQESRIEIPYFECKAF
 PHIFRDLVTCGCRNALRVLPQLQPGNCILLSVSLVRPHIDISFHLGSLINTVF
 VLRSKGLLDVSECEDELTEISCLRLKGMILYLPDASLILFCSPSVNMDDLT
 RGLYLSLDLPHDADLVLLGQFEERYKLTQELVLDRLQLTLRALEDEKKTDT
 LYLVLPPSVANLRLKRPVKKRYDNVTILFSGIVGFNAFCSHKASGGAMKIVNLL
 NDLYTRFDLTDSRKNPFVKYVTGDKYMTVSGLPCECIHHARSICHLADMMETAG
 QVQVDCESQIITIGITGVVTVGIVGORMPRYCLFGNTVNLASRTETGKGNVSE
 YYRCLMTFNSDPQPHLEHRGVPVSMKKKPMQWFLSRKNVTGTEETQDEN"

BASE COUNT 852 a 692 c 723 g 870 t
 ORIGIN
 Query Match 72.6%; Score 1774.2; DB 4; Length 3137;
 Best Local Similarity 85.9%; Pred. No. 0;
 Matches 2036; Conservative 0; Mismatches 308; Indels 27; Gaps 5;

QY 75 CCGGTGCAGACACCATGTAGCGATTTGTGAATCACGCCCTGGAGTTGCTGGTATCGCA 134
 Db 73 CCGGCGCTGACACCATGTATGGATTCTGTAATCACGGCTGGAGTCTGGTATCGCA 132
 QY 135 ATTACGCCGCCAGGTGTGGGAAGACATCAAAAAGAGGCACAGTGTAGATGAAGAAGGAC 194
 Db 133 ACTACGCCCTGAGGTGTGGGAAGACATCAAAAAGAGGCACAAATAGATGAAGAAGGAC 192
 QY 195 AGTTTCTGTGCAGAAATAATATATGATGACTCCAAAATATATGTTGTTGCTGCTGCAA 254
 Db 193 AGTTTCTGTGCCGAATAATTTATGATGATTCTAAAACATATATGTTGTTGCTGCTCAA 252
 QY 255 GCAAGTCTCTCAATCTCAATGCTGGAGAAATCTCCAAATGTTTGGGAAGATGTTTTCG 314
 Db 253 GCAAGTCTCTCAATCTCAAGCTGGAGAAATCTCCAAATGTTTGGGAAGATGTTTTCG 312
 QY 315 TCTTTTGCAGAAATCTGTTTATGATACAAATCTTGGTGTCTGGGCTCTAAATGTCAAG 374
 Db 313 TCTTCTGTCAAGAATCTGTTTATGATACAAATCTTGGGCTCTGGGCTCTAAAGTCAAG 372
 QY 375 AATTTCTACAGAACCTTGATGCTCTCAGCACCACCTTGCTACCATCTACCCAGGATGC 434
 Db 373 AATTTCTCAGAACCTTGATGCTCTCAGCACCACCTTGCTACCATCTACCCAGGATGC 432
 QY 435 GTCACCTTCTCTTTAGTGTGACTGATGCAGAAAGGGCAAGGAGCTCATTTTGCACACT 494
 Db 433 GGGCACCTTCTCTCCGCTGCTGATGGGACAAAGGCAAGGAGCTCATCTGCACTACT 492
 QY 495 ACTCAGAGAGAGAGACTTCAGGATATTGTCTATTGGAATCATCAAAACAGTGGGCAAC 554
 Db 493 ACTCAGAGAGAGAGAGGCTGCAGGATATTGTCTATCGGAATCATCAAGACAGTAGTCA 552
 QY 555 AATTCATGCACTGAAATAGACATCAAGGTTATTTCAGCAAAAGAAATGAAGATGTGATC 614
 Db 553 AGATACATGGCACTGAAATGACATGAAGGTTATTCAACAAAGAAATGAAGATGTGATC 612
 QY 615 ATACTCAATTTTAAATTGAAGAAAAAGAGTCAAAAGAAAGAGATTTTATGAAGATCTTG 674
 Db 613 ATACTCAATTTTGAATTGAAGAAAAAGAGTCAAAAGAAAGAGATTTTATGAAGATCTTG 672
 QY 675 ACAGATTGAAGAAAAATGTTACCCAGGAATCACCGCATCACCCCATATACATTCTGCAAG 734
 Db 673 ACAGATTGAAGAAAAACGGTACCCAGGAATCACCGCATCACCCCATATACATTCTGCAAG 732
 QY 735 CTTTTCCTTTTCATATAATTTTACCGGACCTAGTGTGCTCAGTGTGGCAATGCTTA 794
 Db 733 CTTTTCCTTTTCATATAATTTTACCGGAGGATCTTGTGGTCACTCAGTGTGGCAAGCCTTA 792
 QY 795 TATACAGAGTCTTCCCGGAGCTCCAGGCTGGGAATTCAGGCTCTCTCTGTCTTCTCGC 854
 Db 793 TCTACAGAGTCTTCCCGGAGCTCCAGGCTGGGAATTCAGGCTCTCTCTGTCTTCTCGC 852
 QY 855 TGGTTCCTCTCATATATTGATATTAGTTTCCATGGGATTCCTTCTCATACATCAATACATGTTT 914
 Db 853 TGGTTCCTCTCATATATTGATATTAGTTTCCATGGAAATCTCTCATACATCAATACATGTTT 912
 QY 915 TTGATTGTGAAGCAAGGAAGGATTTGTTGATGTGGAGAAATFAGAAATGTGAGAGTGAAC 974
 Db 913 TCGTGTGTGAAGCAAGGAAGGTTTGTGGATGTAGAGAAATCAGAGTGTGAGAGTGAAC 972
 QY 975 TGACTGGGACTGAGATCAGCTGCTTACGTCTCAAGGGTCAAAATGATCTACTTACCTGAAG 1034
 Db 973 TGACGGGCACTGAAATCAGCTGCTTACGGTCAAGGGTCAAGATGATCTACTTACCTGAAG 1032
 QY 1035 CAGATAGCATACTTTTCTATGTTTCAACCAAGTGTCAACCTGGAGGATTTTCAACGAAG 1094
 Db 1035 CAGATAGCATACTTTTCTATGTTTCAACCAAGTGTCAACCTGGAGGATTTTCAACGAAG 1094


```

Db      1033 CAGACAGCATACTTTCTGCTGCTCACCAGTGTGATGAACCTGGACGACCTGACCCGCG 1092
QY      1095 GAGGCGCTATCTAAGTGACATCCCTCTGCTGATGATGCCACGCGGCGATCTGTTCTTTTGG 1154
Db      1093 GAGGCGCTGACCTGAGTGACATCCCTCTGACGATGCCACCTCGGGACCTCGTGTCTCTGG 1152
QY      1155 GAGAACATTTAGAGAGAAATACAACTCACCCAGAACTGGAATCTCTCACTGACAGCG 1214
Db      1153 GAGAGCAGTTTCAGAGAGAGTACAAGCTGACCCAGGAACCTGGAATCTCTCAAGACAGCG 1212
QY      1215 TACAGCTCAGCTTAAAGAGCCCTGGAGAGATGAAGAAAGAAAGACACACATTCGTGTATT 1274
Db      1213 TGCAGCTCAGTTAGAGCCCTGGAGATGAAGAAAGAAAGACGACAGCTTGTCTGTATT 1272
QY      1275 CTGTCCTCTCCCGCTCTGTGCAATGAGCTGCGGCACAAGCGTCCAGTGGCTGCCAAAA 1334
Db      1273 CTGTCTCTCTCCCTCACTCTGTGGCCACAGAGCTGAGGCACAAGCGTCCGCTGCGTGAAGA 1332
QY      1335 GATATGACATCTGACCATCTCTTTAGTGGCATTTGTGGCTTCAATGCTTCTCTGTAGCA 1394
Db      1333 GATATGACATCTGACCATCTCTCTTACGCGGCAATTTGTGGCTTCAATGCTTCTCTGACGA 1392
QY      1395 AGCATGCACTGGAGAGGAGCCATGAAGATGCTCAACCTCTCTCAACGACCTCTTACACCA 1454
Db      1393 AGCAGCGCTCTGGGAAGGGCCATGAAGATGCTCAACCTCTCTCAATGACCTCTTACACTA 1452
QY      1455 GATTTGACACACTGACTGATTCCTCCGGAAGAAACCCATTTGTTTATAGGTGGAGACTGTG 1514
Db      1453 GATTTGACACTCTGACGAGCTCCCGGAAATCCATTTGTTTATAGGTGGAGACTGTG 1512
QY      1515 GTGACAATATATGACATGAGTGGTTTACCAGAGCCATGCATTCACCATGACCATCCA 1574
Db      1513 GGGACAAGTACATGACATGAGCGGCTGCGGAGCCATGCATTCACCATGACCATCCA 1572
QY      1575 TCTGCCACTCGCCCTTGACATGATGGAATGCTGGCCAGGTTCACAGTAGATGTGTAAT 1634
Db      1573 TCTGCCACTCGCCCTTGATATGATGAGATGCTGGCCAGGTTCACAGTAGATGTGTAAT 1632
QY      1635 CTGTTTCAGATACAAATAGGATACACACTGGAGAGGTAGTTACAGGTGTCATPAGACAGC 1694
Db      1633 CTGTTTCAGATACGATCGGATCCACACAGGGGAGGTAGTACCGGAGTTCATPAGACAGA 1692
QY      1695 GGATCCCTCGATCTCTTTTGGGAATCTCTCAACCTCAGACGCGGACAGAAACCA 1754
Db      1693 GGATCCCTCGATCTCTTTTGGGAATCTCTCAACCTCAGACGCGGACAGAAACCA 1752
QY      1755 CAGGAGAAAGGGAATAAATGCTCTGAATATACATACAGATCTCTTATGCTCCAG 1814
Db      1753 CTGGAGAAAGGGAATAAATGCTCTGAATATACATACAGATCTCTATGACACACAG 1812
QY      1815 AAAATTCAGATCCCAATTCCTTTGGAGACAGAGCGCCAGTGTCCATGAAGGCAAAA 1874
Db      1813 AAAATTCAGACCGCAATTCCTTTGGAGACAGAGCGCCAGTGTCCATGAAGGCAAAA 1872
QY      1875 AAGAACCAATGCAAGTGTGTTCTTATCCAGAAAAATACAGGAACAGAGGAAACAAAGC 1934
Db      1873 AGGACCCAAAGTGTGTTCTTATCCAGAAAAATACAGGAACAGAGGAAACAAAGC 1932
QY      1935 AGGATGATGACTGAATCTTGGATTATGGGGTGAAGAGGATACAGACTAGGTTCAGATTT 1994
Db      1933 AGGATGAAATGAAATCTGACTGTGGGTTAAGA----TCGCAGGTGAGGTGAGTCT 1988
QY      1995 TCTCCTAACAGCTGCCAGCCAGGAGAGTCTTCCCTATGGATACAGATTTTCTTTTG 2054
Db      1989 ACCCTTAACAGATGCATGCCAGGACCAAGTATTCCTCCCTCCCTCCG-----T 2031
QY      2055 TCTTGTCTCATTTACCCCAAGCTTTCTCTAGATATATCTCTCACTATCCGTTATTCAC 2114
Db      2032 CCTTTTCATGTTTCCCAAGCCCTTCTCTGATGATCTTCTCTTATTCATTCATTCAGC 2091
QY      2115 CTAGCTCTGCTTTCTTACTACTTTTAGCTTTAGTATATATCTAAAGTTTGGCTTTTG 2174
Db      2092 TTAGCTCTGCTTTTGAATATTTTGGAGTTTAGTATGTTATCTGAAGTTTGGCTTT-- 2149

```

```

QY      2175 ATGTGATGATGTGAGCTTTCATGTGCTTTAAATCTACTACAGCATTTACCTAAACATGCT 2234
Db      2150 GTGTGGCTGATGTGAGCTTTCACATGCTTTCAGATCGACTAC--GCATCGCTGGCATGCT 2207
QY      2235 GATCTGCAAGTAGTAGGACCCCAATTAATATTATTGTTGAATTTAGTTAAATGAACACTGAAC 2294
Db      2208 GGTTCCTCAAGTAGTGGGACCCCAATTAATATTATTGTTGAATTTAAATGAACACTGAAC 2267
QY      2295 AGTGTTCGCCATGTGTATATTATATATATATATATATATATATATATATATATATATATAT 2352
Db      2268 AGTATTCAGCCATGTGTATATATATATATATATATATATATATATATATATATATATATAT 2327
QY      2353 TATATGTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2412
Db      2328 TCTCTATATGTGTATATATATATATATATATATATATATATATATATATATATATATATATAT 2387
QY      2413 TATCATATATAGAAATCAATTTCTTAAAGGAGT 2443
Db      2388 TATCATATATAGAAATCGTTTTTCTTAAAGGAGT 2418

RESULT 5
AF297083 3235 bp mRNA linear ROD 01-OCT-2000
LOCUS Mus musculus soluble guanylyl cyclase beta 1 subunit mRNA, complete
DEFINITION cds.
ACCESSION AF297083
VERSION AF297083.1 GI:10442715
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3235)
AUTHORS Sharina, I.G., Krumenacker, J.S., Martin, E. and Murad, F.
TITLE genomic organization of alpha and beta subunits of the mammalian soluble guanylyl cyclase genes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (20), 10878-10883 (2000)
MEDLINE 20461464
PUBMED 10984516
REFERENCE 2 (bases 1 to 3235)
AUTHORS Krumenacker, J.S., Sharina, I.G., Martin, E.S. and Murad, F.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2000) Integrative Biology and Pharmacology, University of Texas Medical School, 6431 Fannin Street, Houston, TX 77030, USA

FEATURES
source
1..3235
/organism="Mus musculus"
/strain="C57BL/10"
/db_xref="taxon:10090"
112..1974
/codon_start=1
/product="soluble guanylyl cyclase beta 1 subunit"
/protein_id="AAG17447.1"
/db_xref="GI:10442716"
/translat="MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEQFLVRIYD
DSKTDLVAAASKVLNLAGETILQFMKFFVCOESGYDTILVLGNSVRFQNLID
ALHDHLATYPCMRAPSPRCTDAEKKGILHLYSERGLQDIVIGIKTVQALHGT
EIDMKVIQORNEECDTQFLIEKESKEDEYDDDRFEENGTSQSRISPTFCRAFP
FHIIDRLNVTCGNALYRLVLPQPGNSLLSVFLPHIDISFHLGSHINTVF
VLRSKGLDVEKLECEBELTGAEISCLRLKQMYLPEADSLFICSPSVNLDLDT
RRGLYLDLIDHATDVLVLLGEFREYKLTQELIETDLRLQLTARALEDKKTDIT
LLY SVLPYVANELRHKRPVPAKRVNVTILFGIVGPNACFCGSHASGKAMKVNLL
NDLYTREFDITD SRKNPFVYKVETVGDKMYTVSGLPFPCIHARSICHLALDMEIAG
QVQVGESQITIGIHTEVTVGICQMRPCLFNTGNTGKPKMPQVFLSRKNTGTEETNEEN"
BASE COUNT 901 a 673 c 741 g 920 t
ORIGIN
Query Match 70.1%; Score 1711.4; DB 10; Length 3235;
Best Local Similarity 83.7%; Pred. No. 0;

```

Matches	2052;	Conservative	0;	Mismatches	376;	Indels	25;	Gaps	9;
---------	-------	--------------	----	------------	------	--------	-----	------	----

QY	9	CGCCGTCGCCCTCTCTCCTGGGTCCCTTCGCGCGTACCTCTCGTGGGGGCTGCCCTCCC	68
Db	32	CGCGCGTGTAGTGCCTGTCTGGGTCCCTTCGCGGGTACTCTGCCTGGAAGCTGCTTCC	91
QY	69	CGGCTCCGGTCGAGACACCATCTACGGATTCTTGATCATCGCCCTGGAGTTGCTGTGTA	128
Db	92	AGGCTCCGGCGGGACACCATCTACGGTTCTGTGAACCATGCGCTCGAGCTGCTGTGTA	151
QY	129	TCGCAATTAACGCCCGCAGGCTGTGGGAACACATCAAAAAAGGACACAGTTAGATGAAG	188
Db	152	TCCGCAATTATGTCGCCAGGTCGTGGAACACATCAAAAAGAGGCACACCTGGATGAAG	211
QY	189	AAGACAGTTCTTGTGCAGAAATATATGATGACTCCAAAACCTATATGATTTGGTTGCTG	248
Db	212	AAGCCAGTTCTTGTGCAGAAATATATACATGATTCCAAAACATATGACTTGGTGGCTG	271
QY	249	CTCAAGCAAGTCCCTCAATCTCAATGCTGGAGAAATCCTCCAAATGTTTGGGAAGATGT	308
Db	272	CTCCAGCAAGTCTCAACCTCAATGCTGGCGAAATCCTGCAAGATTTTGGGAAGATGT	331
QY	309	TTTTTCGTTTTCGCCAAGAACTCGGTTATGATACAATCTTTCGCTGTCTCGGCTCTAATG	368
Db	332	TTTTTCGTTTTCGCCAGGAGCTCGGCTATGATACCATCTTTCGCTGTCTCGGATCTAATG	391
QY	369	TCAGAGAAATTCCTACAGAACCTTGTGCTCTGCAGCACCACTTGTCTACCATCTACCCAG	428
Db	392	TCAGAGATTTTTCAGAACCTCGATGCCCTGATGACCACTTCGCCACCATTTACCCAG	451
QY	429	GAATGCTGCACCTTCTCTTTAGTGTGACTGATGCAGAAAGGCGAAAGGACTCATTTTGC	488
Db	452	GGATGCGCGCGCTTCTCTTCAGGTGCACGATCGGGAGAAAGGCAAAAGGCTCATCTCTG	511
QY	489	ACTACTACTCAGAGAGAGAGGACTTCAGATATTTGTCATTTGGNATCATCAAAACACGTG	548
Db	512	ACTACTACTCGGAAGAGAGGGGCTTCAGACATCTGTCGATCGGGATTTATCAAGACTCTT	571
QY	549	CACAACAAATCCATGGCACTGAAATGACATGAAGGTTATTCAGCAAGAAATGAAGAAT	608
Db	572	CTCAACAGATACACGCACTGAGATAGACATGAAGGTGATTCAGCAAGAAATGAAGAAT	631
QY	609	GTGATCATACTCAATTTTAAATTTGAAGAAAGACATCAAAAGAGAGGATTTTATGAAG	668
Db	632	GTGATCACACCCAGTTTAAATTTGAAGAAAGGAATCAAAAGAGAGGATTTTATGAAG	691
QY	669	ATCTTCACAGATTTGAAGAAATGTTACCCAGAAATCACGCATCAGGCCATATACATCTCT	728
Db	692	ATCTCGATAGTTTGAAGAAATGTTACCCAGAAATCACGCATCAGGCCATACCTTCT	751
QY	729	GCAAGCTTTTCTTTTCATATAATATTTGACCGGACCTTAGTGGTCACATCAGTGTGGCA	788
Db	752	GCAAGCATTTCTTTTCATCATATATTTGACCGGAACCTTAGTGGTCACATCAGTGTGGCA	811
QY	789	ATGCTATATACAGATTTCTCCCGCTCCAGCTGGGAATTCAGCGCTTCTGCTCTCTCT	848
Db	812	ATGCCATCTACAGATGCTCCCGCTCCAGCTGGGAATTCAGCGCTTCTGCTCTCTCT	871
QY	849	TCCTCGTTGTTCTCTCATATTTGATATTTAGTTTCCATGGGATCTTCTTCACATCAATA	908
Db	872	TCCTCTGGTCCGCCCTTCATCGACATCAGTTTCCATGGGATTTCTTCACATCAATA	931
QY	909	CTGTTTTTTGATTTGAAAGCAAGAGGATTTGTTGGATGTGGAGAAATTAAGATGTGAGG	968
Db	932	CAGTCTTTGTACTTGAAAGCAAGAGGGTTGTTGGATGTGGAGAAATTTAGTGTGAGG	991
QY	969	ATGAATGATGGGACTGAGATCAGCTGCTTACGCTCAAGGGTCAAAATGATCTACTTAC	1028
Db	992	ATGAATGATGGGACGAGATTTAGTGTCTAGTCTCAAGGCCAATGATCTACTTAC	1051
QY	1029	CTGAAGCAGATACGATCTTTTCTATGTTTCACCAAGTGTGATGAACCTGGACGATTTGA	1088
Db	1052	CAGAAGCAGATACGATCTTCTCTGTTTCAACAAGTGTGATGAACCTGGAGGACCTAA	1111

Qy	1089	CAAGGAGAGGCGCTGTATCTAAGTGTACATCCCTCTGCATGATGCCAGCGGGATCTTGTGTTCT	1148
Db	1112	CAAGAAGAGCCCTGTATCTGAGTGACATCCCTCTCCAGCATGTCTACCCGAGACCTGTGTTCT	1171
Qy	1149	TTTTGGGAGAAACAATTTAGAGAGGAATACAAACTCACCCAAAGAACTGGAAATCCTCACATG	1208
Db	1172	TTTTGGGAGAAACAGTTCCCGGAGGAGGTACAAACTGCACAAAGAGCTGGAAATCCTCACCG	1231
Qy	1209	ACAGGCTACAGCTCACGCTTAAGAGCCCTTGAAGATGAAAGAAAAAGACAGACACATTCG	1268
Db	1232	ACAGGCTGCAGCTCACACTGAGAGCCCTTGGAGGATGAGAAGAAAAAGACAGACACATTCG	1291
Qy	1269	TGTATTCTGTCCCTTCCCTCCGCTCTGTGGCCAAATGAGCTGGGGCACAAGCGTCCAGTGCCTG	1328
Db	1292	TGTATTCTGTCCCTTCCCTCCATCTGTGGCCAAATGAGCTGAGACAAGACGCCAGTGCCTG	1351
Qy	1329	CCAAAAGATGACAATGTGACCATCCTCTTTAGTGGCAATGTGGGCTTCAATGCTTTCT	1388
Db	1352	CCAAAAGATGACAATGTGACCATCCTCTTACGGCGCATTTGTGGCTTCAATGCTTTCT	1411
Qy	1389	GTAGCAAGCATGCAATCTGGAGAAGAGGCCAATGAAGATGCTCAACCTTCCTCAACGACCTCT	1448
Db	1412	GTAGCAAGCATGCAATCTGGAGAAGAGGCCAATGAAGATGCTCAATCTCCTCAACGATCTCT	1471
Qy	1449	ACACCGATTTGACACACTGACACTGACTTTCCCGGAAAAACCCATTTGTTTAAAGTGGAGA	1508
Db	1472	ACACCGATTTGACACACTGACACTGACTTTACGAAAAAACCCATTTGTTTACAGAGTGGAAA	1531
Qy	1509	CTGTTGGTGACAAGTATATGACAGTGAAGTGGTTTACCAGAGCCATGCATTCACCATGCAC	1568
Db	1532	CAGTTGGTGACAAGTATATGACAGTGAAGTGGCTTGGCAGAACCTTGTATCCACCATGCAC	1591
Qy	1569	GATCCATCTGCCACCTGGCCCTGCGACATGATGAAATGCTGGCCAGGTTCAAGTAGATG	1628
Db	1592	GGTCCATTTGCCACCTGGCTTTAGACATGATGAAATGCTGGTCAAGTTCAAGTAGATG	1651
Qy	1629	GTGAATCTGTTCCAGATAACAATAGGATACACACTGGAGAGGTAGTTACAGGTGTCTATAG	1688
Db	1652	GTGAATCTGTTCCAGATAACAATCGGATCCATACCGGGAGGTTGGTACAGGTGTCTATG	1711
Qy	1689	GACAGCGGATGCCCTCGATCTGCTTTTGGGAATACTGTCACCTTCACNAGCCGAAACAG	1748
Db	1712	GACAGCGGATGCCCTCGGATTTGCTCTTTGGGAATACCGGTCAACCTTCACNAGCAGACAG	1771
Qy	1749	AAACCACAGAGAAAAAGGGAAAAATAAATGTGCTGAATATACATACAGATGCTCTATGT	1808
Db	1772	AAACCACAGAGAAAAAGGGAAAAAGATAAATGTTCCGAAATATACATACAGGTGTCTATGT	1831
Qy	1809	CTCCAGAAATTCAGATCCAAATTCACATTTGGAGCACAGAGGCCAGTGTCCATGAAGG	1868
Db	1832	CTCCAGAAATTCAGATCCAAATTCACATTTGGAGCACAGAGGCCAGTGTCTATGAAGG	1891
Qy	1869	GCAAAAAGAACCAATGCAAGTTTGTCTTCTATCCAGAAAAAATACAGGAACAGAGGAAA	1928
Db	1892	GCAAGAGGAACCAATGCAAGTCTGTTCTCTATCCAGAAAAAATACAGGCACGGAGGAAA	1951
Qy	1929	CAAAAGCAGGATGATGACTGAATCTTGATTTATGGGTTGAAGAGGAGTAC--AGACATAGGT	1986
Db	1952	CAAAATGAGGAGGATGAAACTGAGTGTGAGTTTGGAAACAAGAGAACGTGCTGTTGGGT	2011
Qy	1987	TCCAGTTTTCTCTTAACAGTGGCAAGCCGAGGAGTCTTCCCTATGGAATACAGATT	2046
Db	2012	TCGGGTGACAGTCTAATGTGTGTCAAAGCAGGAGCACTTCTTTCCCTGTGGATPAGCAAT	2071
Qy	2047	TTCTTTTGTCTTGTCCATTAACCCCAAGACTTTCTTCTAG-----ATATATCTCTCACTAT	2102
Db	2072	T--CTACTTCGTCTTTAGTGGCCCAAGGCTTTCCTTAGTTACACAGATCTCACACTAT	2129
Qy	2103	CCGTTATTAACCTTAGCTCTGCTTCTTAATTAATTTTTAGGCTTTAGTATATTAATCTAAA	2162
Db	2130	GGTTTATTTTATGATTTTATGCTCTGCTTCGATTAATTTTTTAAAGTCTCAGTATATTTCCAAA	2189

Qy	2163	GTTT-GCCTTTTGTGATGGGAATGATGAGCTTCATGTGCTTAAATCTACTACAAGCAT	2222
Db	2190	TTTTGGGTTTTTGTGATGGGACTGTTGAGC- ---TGTTTTCTAAATTTCTCTACAAGCAT	2245
Qy	2222	TACCTAACATGGTGATCTGCAAGTAGTAGGCCACCCCAAT ---AAATATTTGTTGAATTTAG	2278
Db	2246	TGCTTAGACGGAGCTTCTAAAGTGATGAAGCACCCACCTGTTAAAGTTTGTAAATCTGA	2305
Qy	2279	TTAAATGAACATGAACAGAGTGTGGGCATGTGTATATTTATATCATGTTTTACCAAAATCTG	2338
Db	2306	TAGAAGCAGACTTAATAGTAGTATCGGCCATCGGTGTATATAT-CATGGCTCAGTAGATTG	2364
Qy	2339	TTTAGTGTGCCACATATA-----TGTATATGATATATTTTAATGACATATA--ATGTAATA	2390
Db	2365	TTTTATGCTCCATGTATATGTGTGTATATGATATATTTAAATGACTATATACCATAAAACA	2424
Qy	2391	AAGTTTATATCATGTGGTGTATATCATATATAGAAATCATTTCTAAAGAGT	2443
Db	2425	AAGTTTATATCATGTGGTGCATGGCATCTTAGAACCAATTTGTACAGAGT	2477
RESULT 6			
RATGCS		3047 bp mRNA linear ROD 27-APR-1993	
LOCUS		Rat soluble guanylate cyclase 70kd subunit mRNA, complete cds.	
DEFINITION		M22562	
ACCESSION		M22562.1 GI:204273	
VERSION			
KEYWORDS		guanylate cyclase.	
SOURCE		Rat (Wistar; adult) lung, cDNA to mRNA, clones 70-36, 36-8, 36-16.	
ORGANISM		Rattus norvegicus	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
REFERENCE		1. (bases 1 to 3047)	
AUTHORS		Nakane, M., Saheki, S., Kuno, T., Ishii, K. and Murad, F.	
TITLE		Molecular cloning of a cDNA coding for 70 kilodalton subunit of soluble guanylate cyclase from rat lung	
JOURNAL		Biochem. Biophys. Res. Commun. 157 (3), 1139-1147 (1988)	
MEDLINE		89087429	
PUBMED		2905128	
COMMENT		Draft entry and computer-readable sequence for [1] kindly provided by M.Nakane, 15-FEB-1989.	
FEATURES			
source		Location/Qualifiers	
		1..3047	
mRNA		/organism="Rattus norvegicus"	
		/db_xref="taxon:10116"	
		<1..3047	
CDS		/product="SGC mRNA"	
		42..1901	
		/codon_start=1	
		/protein_id="AAA41204.1"	
		/db_xref="GI:204274"	
		/translation="MYGFVNHAELELVIRNYPVEWEDIKKEAQLDQEGQFLVRIYDHSYDILVAASVKNLNAAGETLQMEGKMFVFCQSEGYDILRLVGSNREFQLNDALHDLYATIPGMRAFSFRCDEAGKGLILHRYSEGYQDVIYGIKTTVAQIHGT EIDMKYIQRSEEDCTHPILESKESEDFEDLDLRFENGQDSRISPYTCKAFPI FHIIFDRILVTCGNAIVYLPQEGKCSLLSVFLSRPHDIDISPHGLTSLHNVPV VLRSKGLDVLKECEDELGTAEISCLRLKGOMIYLPADSLFCLSPSPVMNLDLTL RRGILSLPDLPATRDLDVLGGQFKEEYKLTLEILLDRLQLTLKALDEKKYDILL YLSVLPSPVANELRHRPYPARYDNVTILFSGIVGFNAFCSHKASGAGKMYVALL NLDYRFDITLDRKNPFVYKVTGVDGKYMTVSLGPECIHHARSICHLALDMMETAG QVQDQESVQITTGITGVEGVIGORMPRYCLFGNTVNLTRTETTTGPKGKINYS YTRCLMSPENSQDPLEHRRGPVSMKGKPEQVQVFLSRKNTGTEETNQDEN"	
BASE COUNT		850 a 656 c 691 g 850 t	
ORIGIN		244 bp upstream of psti site.	
Query Match		69.3%; Score 1691.8; DB 10; Length 3047;	
Best Local Similarity		83.6%; Pred. No. 0;	
Matches 2019;		Conservative 0; Mismatches 372; Indels 24; Gaps 8;	
Qy	50	TGCGTGGGGCTGCCTCCCGGCTCCCGGTGCACACACCATGTACGGATTGTGAATCAC	109

Db 32 GCGCGCTGTAGTCGCTGTCTGGGTCCCTTCGCGGCTACTCTTGCTGGAACTGCTTCCT 91
Qy 69 CGGCTCCGCGTGCAGACACCATGATGACGATTTGTGAATCACGCCCTGGAGTGTCTGTGA 128
Db 92 AGGCTCCGCGCGGCGACACCATGATGACGATTTGTGCGTGAACCATGCTCGAGCTGTGTGA 151
Qy 129 TCCGCAATTAGCGCGCGAGTGTGGAGACATCAAAAAGAGGCACAGTATGATGAAG 188
Db 152 TCCGCAATATGCGCGAGTGTGGAGACATCAAAAAGAGGCACAGTGTGATGAAG 211
Qy 189 AAGGACAGTTCCTGTGAGATAATATATGATGACTCCAAATTTATGATTTGGTGTGCTG 248
Db 212 AAGCCAGTTCCTGTGAGATAATATGATGATGATTCAAAACATATGACCTGTGTGCTG 271
Qy 249 CTGCAAGCAAGTCTCAATCTCAATGCTGAGAAATCTCCAAATGTTGGGAGATGT 308
Db 272 CTGCAAGCAAGTCTCAATCTCAATGCTGCGAAATCTCGAGATGTTGGAGATGT 331
Qy 309 TTTTCGCTTTTGCAGAACTGTTATGATACAACTTTGCGTGTCTGGCTCTAATG 368
Db 332 TTTTCGCTTTTGCAGAGTCTGGCTATGATACCATCTTGGTGTCTGGGATCTAATG 391
Qy 369 TCAGAGATTTCTACAGAACTTGATGCTCTGCAGCACCATCTTGTACCATCTACCCAG 428
Db 392 TCAGAGATTTTGCAGAACTCGATGCCCTGATGACCATCTGCACCATTTACCCAG 451
Qy 429 GAATGCGTGCACCTTCTTTCAGTGTGCTGATGAGAAAGGGCAAGGACTCAATTTTCG 488
Db 452 GGATGCGCGCTTCTTTCAGTGTGACCATGATGCGGAGAAAGGCAAGGCTCATCTTCG 511
Qy 489 ACTACTCTAGAGAGAGAGGACTTCAGGATATGTCATTTGGATCATCAAAACAGTGG 548
Db 512 ACTACTCTAGAGAGAGAGGCTTCAGGATATGTCATTTGGATCATCAAAACAGTGG 571
Qy 549 CACAAATTCATGCGACTGAAATAGACATGAAGTTTATTCAGCAAAAGAAATGAAGAT 608
Db 572 CTCACAGATACAGGCACTGATGATGACATGAAGTGTTCAGCAAAAGAAATGAAGAT 631
Qy 609 GTGATCATCAATTTTAAATGAAGAAAGAGTCAAAAGAGAGGATTTTATGAAG 668
Db 632 GTGATCACACCCAGTTTAAATGAAGAAAGAGTCAAAAGAGAGGATTTTATGAAG 691
Qy 669 ATCTTGACAGATTTGAAGAAATGTCACCCAGGATACGATCAGCCCATATACATCT 728
Db 692 ATCTGGATAGTGTGAAGAGATGTACCCAGGATACGATCAGCCCTTACACCTTCT 751
Qy 729 GCAAGCTTTTCCCTTTTCAATATATTTGACCGGCACTAGTGTCTACTCAGTGTGCA 788
Db 752 GCAAGCATTTCCCTTTTCAATATATTTGACCGGCACTAGTGTCTACTCAGTGTGCA 811
Qy 789 ATGCTATATACAGATTTCCCTCCAGCTCCAGCTGGGAATGCAAGCTTCTGTCTGTCT 848
Db 812 ATGCCATCTACAGATGTCTCCCTCCAGCTCCAGCTGGGAATGCAAGCTTCTGTCTGTCT 871
Qy 849 TCTGCTGTGTCTCTCATATGATATTTGATTTCCATGGGATCTTCTCACATCAATA 908
Db 872 TCTCTGTGTCTCTCATATGATATTTGATTTCCATGGGATCTTCTCACATCAATA 931
Qy 909 CTGTTTGTGTATGAGAGCAAGGAGGATTTGATGATGAGAGAAATTTAGAAATGTGAGG 968
Db 932 CAGTCTTTGTACTGAGAGCAAGGAGGTTGCTGATGTGGAGAACTTGTAGTGTGAGG 991
Qy 969 ATGAAGTGTGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028
Db 992 ATGAAGTGTGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1051
Qy 1029 CTGAAGCAGATAGCATATTTTCTATGTTACCAAGTGTGATGATGATGATGATGATGAT 1088
Db 1052 CAGAAGCAGATAGCATATTTTCTATGTTACCAAGTGTGATGATGATGATGATGATGAT 1111
Qy 1089 CAAGAGAGGGTGTATCTAAGTGACATCCCTCTGCATGATGATGATGATGATGATGAT 1148
Db 1112 CAAGAGAGGGTGTATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1171

Qy 1149 TTTTGGGAGAACAAATTTAGAGAGGAATACAAACTCACCCAAAGAACTGGAAATCTCTACTG 1208
Db 1172 TTTTGGGAGAACAGTTCCGGGAGGAGTACAAACTCACCAAGAGCTGGAAATCTCTACCG 1231
Qy 1209 ACAGGCTACAGCTCAGTTAAGAGCCCTTGAAGAGCTTGAAGATGAAGAAAGAGAGACATGCG 1268
Db 1232 ACAGGCTCAGCTCAGCTGAGAGCCTTGGAGGATGAGAAAGAAAGAGACATGCG 1291
Qy 1269 TGTATTCGTCTCTCTCCCTGCTGTTGCAATGAGCTGCGGCAAGGCTCCAGTGCCTG 1328
Db 1292 TGTATTCGTCT 1351
Qy 1329 CCAAAAGATATGACAAATGTGACCATCTCTTTAGTGGCAATGTGGGCTTCAATGTTCT 1388
Db 1352 CCAAAAGATATGACAAATGTGACCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1411
Qy 1389 GTAGCAAGCATGCTCTGAGAGGAGCATGAAGATGTCACCTCTCTCAAGGACCTCT 1448
Db 1412 GTAGCAAGCATGCTCTGAGAGGAGCATGAAGATGTCACCTCTCTCAAGGACCTCT 1471
Qy 1449 ACACGAGATTTGACACACTGACTGATTCGCGGAAAGAAACCATTTGTTTATAAGTGGAGA 1508
Db 1472 ACACCGATTTGACACACTGACTGATTCAGGAAAGAAACCATTTGTTTACAGGTTGAAA 1531
Qy 1509 CTGTTGGTGACAAATGATGACAGTGTGAGTGTGTACAGAGCCATGATTCACCATGCGAC 1568
Db 1532 CAGTTGGTGACAAATGATGACAGTGTGAGTGTGTGAGAGCCATGATTCACCATGCGAC 1591
Qy 1569 GATCCATCTGCCACCTGCGCTTGGACATGATGGAATGCTGCGGAGGTTCAAGTAGATG 1628
Db 1592 GGTCCATTTGCCACCTGCGCTTTAGACATGATGGAATGCTGCGTCAAGTGTAGTAGATG 1651
Qy 1629 GTGAATCTGTTGAGATAACAAATGAGGATACACACTGGAGAGGTAGTTACAGGTGTCTATAG 1688
Db 1652 GTGAATCTGTTGAGATAACAAATGAGGATACCAATCGGAGGAGTGTGACAGGTGTCTATG 1711
Qy 1689 GACAGCGGATGCTCGATCTCTCTTTTGGGAATCTGTCACCTCTCAAGCCGCAAGACAG 1748
Db 1712 GACAGCGGATGCTCGGATTTCTCTTTTGGGAATCTGTCACCTCTCAAGCCGCAAGACAG 1771
Qy 1749 AAACACAGGAGAAAGGAAATATAATGCTCTGTAATATACATACAGATGCTTATGCT 1808
Db 1772 AAACACAGGAGAAAGGAAATATAATGCTCTGTAATATACATACAGATGCTTATGCT 1831
Qy 1809 CTCCAGAAATTCAGATCCACAAATTCCTCTTGGAGCAGAGGCCAGTGTCCATGAAGG 1868
Db 1832 CTCCAGAAATTCAGATCCACAAATTCCTCTTGGAGCAGAGGCCAGTGTCTATGAAGG 1891
Qy 1869 GCAAAAGAACCAATGCAAGTTTGGTTCATCCAGAAATAATCAGAGAACAGAGGAAA 1928
Db 1892 GCAAGAGAGAACCAATGCAAGTTTGGTTCATCCAGAAATAATCAGAGAACAGAGGAAA 1951
Qy 1929 CAAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1986
Db 1952 CAAATGAGAGGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2011
Qy 1987 TCCAGTTTCTCTTAACAGCTGCCAAGCCAGGAGCAGTCTTCCCTATGATGATGATGAT 2046
Db 2012 TCGGTTGACAGTCTAATGTTGTCAGCAAGCAAGGAGCAGTCTTCTTCCCTGTGATGATGAT 2071
Qy 2047 TCTTTTGTCTCTGTCATTTACCCCAAGACTTTCTCTCTAG----ATATATCTCTCTCTCT 2102
Db 2072 T--CTACTCTGTTTGTAGTGGCCCAAGGCTTCTCTCTAGTTTACAGATCTCACACTAT 2129
Qy 2103 CGTTTATCAACTTAGCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2162
Db 2130 GGTATTTGATTTTAGCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2189
Qy 2163 GTTT--GGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2221
Db 2190 GTTTGGTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2245

Db 311 AAATGGGAAGAAATCAACAGCGAAGCCGAGTGGAGATCGAAGGCGAGTTTCTTAGTCGA 370
QY 209 ATAATATATGATGACTCAAAACATTATGATTGGTTGCTGCTGCAAGCAAAAGTCTCTCAAT 268
Db 371 ATCGGTATGATGATGCTTATCTATCTATGATGATGATGATGATGATGATGATGATGATGAT 430
QY 269 CTCATGCTGGAGAAATCTCCAAATGTTGGGAAGATGTTTTCGTTCTTGGCCCAAGAA 328
Db 431 ATAAGTGGCAATGATCTGTTGGAGGCAATCGGTGGATGTTCTTTGAATTTTGGCTCGAG 490
QY 329 TCTGTTATGATACAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 388
Db 491 TCTGCTACGACAAACATCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 550
QY 389 CTGATGCTGCGACGACACCTGCTACCAATCTACCAAGGAAATGCGTGGCACTTCTTCTT 448
Db 551 TTAGATGCTTCTTCAAGATCTTCTGCTCAATCTACCGGGAATCGGAGCACTTCAATTC 610
QY 449 AGGTGCACTGATGACAGAAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 508
Db 611 AGATGCACTACAGAGAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670
QY 509 GGACTTCAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568
Db 671 GGGCTGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
QY 569 GAAATAGACATGAAGTTTATGACAAAGAAATGAAGAAATGATGATGATGATGATGATGATGAT 628
Db 731 GAAATGCTGCTGAGATCATCAAGAAATGAAGAAAGAGTGGGATGCTGCTGCTGCTGCTGCT 790
QY 629 ATTGAAGAAAGAGTCAAGAGAGAGGATTTTATGAAGATCTGACAGATTTTGAAGAA 588
Db 791 ATCATGAGAGAGTGGAGACAGGAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 847
QY 689 AATGTTACCCAGGAAATACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
Db 848 GCTCTCTCAAGGACCAAAATCAGTCCATCAACACTGTCGAGGATCCTGCCCTTCCAC 907
QY 749 ATAATATTTAGCCGGGACCTAGTGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808
Db 908 ATCATGTTCAATGCTGTAATTAACCTAGAGCAAGCAGGTAAGTCCATTCAAAGGATGTA 967
QY 809 CCCAGCTCCAGCTGGGAATGACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868
Db 968 CTAACATATCAATCTTCACTGCAAGATGAGGAGTATTCATGATGCTGCTGCTGCTGCTGCT 1027
QY 869 ATTGATATAGTTTCCATGGGATCCTTCTCACATCAATGCTGCTGCTGCTGCTGCTGCTGCTG 928
Db 1028 ATGGAGTTTCACTTCAAGTCTATCTCTCAGTCAATGCAACACACATTTACGCTGCTGAAAC 1087
QY 929 AAGGAGGATGTTGGATGTTGGAGAAATAGATGTAGGATGAATGACTGCTGCTGCTGCTGCT 988
Db 1088 AACTCTGTTGTTAAATCCAAACA-----ACCCCGTAACGGCTCC 1129
QY 989 ATCAGCTGCTTACGCTCAAGGCTCAAAATGATCTACTTACCTGAAGCAGATAGCATCTT 1048
Db 1130 ATCCAGCACTCAAGCTCAAGGCGGATGCTTACGTAACAGATGCAATGATGATGATGATGAT 1189
QY 1049 TTTCTATGTTCAACAGTGTATGAACTTGGACGATTTGACAAAGGAGGCTGCTGCTGCTGCT 1108
Db 1190 TATCTCTGTTCAACACAGTCTATCAATCTTATGAGCTGAGGCAACGTAAGTGTACCTG 1249
QY 1109 AGTGACATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1168
Db 1250 AGTGATATCCATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309
QY 1169 GAGGATCAAACTCAACCAAGTGGAAATCTTCACTGACGCTGACGCTGACGCTGACGCTG 1228
Db 1310 GAGGAGTCAAACTCAACCAAGTGGAAATCTTCACTGACGCTGACGCTGACGCTGACGCTG 1369
QY 1229 AGAGCCCTGGAAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1288
Db 1370 CTTGAGATCGAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1429

QY 1289 TCTGTTGCAATGAGTGGCGCAAGCGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1348
Db 1430 TCCGTTGCTAAGAGTTAGTCTATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1489
QY 1349 ACCATCTCTTCTAGTGGCATTTGTTGGCTTCAATGCTTCTGTTAGCAAGCATGCTGCTG 1408
Db 1490 ACTCTGATGTTTCAAGTGGATCTTTGGCTTTGAGAGCTTTTGT-----CGTGGCTACTCA 1543
QY 1409 GAAGGAGCATGAAGATGCTCAACCTCCTCAACGACCTCTACACGAGATTTTACACACTG 1468
Db 1544 CACGAGCTATGAAGATGTTAGCTTCTTACAGTCTTACAGTCTTACAGTCTTACAGTCTT 1603
QY 1459 ACTGATCCCGGAAACCCATTTTATTAAGTGGAGCTGTTTGGTGACAAAGTATATG 1528
Db 1604 ATGGA-----AATAACCCAGACGTTTACAAAGTAGAGACAGTGGTGTATTAATACATG 1657
QY 1529 ACAGTGAAGTGGTTTACCAGAGCCATTCATTCCACATGACAGTCCATCTGCGACCTGGCC 1588
Db 1658 GCAGTAAGTGGTCTACCCGTACCATGCTGATCATGCTAAATGATCGCTAAGATGGG 1717
QY 1589 TTGGACATGATGGAATTCCTGGCCAGGTTCAAGTAGATGGTGAATCTGTTCCAGATAACA 1648
Db 1718 CTTGAATGAGGAATCTCTGACATGTCATCATGAGGAGTACCTATAGTATGATCT 1777
QY 1649 ATAGGATACACACTGGAGAGTGTATAGGTGTCATAGGACGCGATGCTCGATAC 1708
Db 1778 ATTGGTCTTACAGTGGAGAGTGTGACAGAGTGGTGGGACAACGATGCGAGGTAC 1837
QY 1709 TGTCTTTTGGGAATCTGTCACCTCACAACCCCAACAGAAACACACAGGAGAAAGGA 1768
Db 1838 TGTCTTTTGGTGAACCTGTCACCTCCAGGACAGAGACTACTGTTGTCACGGC 1897
QY 1769 AAAATAAATGCTCTGAATATACATACAGATGCTTATGCTCCAGAAATTCAGATCA 1828
Db 1898 AGATCAACATCCCGATACCTTACGATGTTTGTATGAGGACGACAGATGCAACCA 1957
QY 1829 CAATTCCTACTGGACACAGAGCCAGTGTCCATGAAGGCGAAAGAAACCAATGCA 1888
Db 1958 ACCTTCCAGTTCGATTTCCGAGGTCTAGTCAACATGAAGGCAAGCTAAACCTGCCCA 2017
QY 1889 GTTGGTCTTCTATCCAGAAA 1909
Db 2018 TGCTATCTGCTCTCAAGAAA 2038

RESULT 11
AF062751
LOCUS Manduca sexta soluble guanylyl cyclase beta-1 subunit mRNA, complete cds.
ACCESSION AF062751
VERSION AF062751.1 GI:3372755
KEYWORDS
SOURCE Manduca sexta.
ORGANISM Manduca sexta
REFERENCE 1 (bases 1 to 3691)
AUTHORS Nighorn,A., Gibson,N.J., Rivers,D.M., Hildebrand,J.G. and Morton,D.B.
TITLE The nitric oxide-cGMP pathway may mediate communication between sensory afferents and projection neurons in the antennal lobe of Manduca sexta
JOURNAL J. Neurosci. 18 (18), 7244-7255 (1998)
MEDLINE 9840973
PUBMED 9736646
REFERENCE 2 (bases 1 to 3691)
AUTHORS Nighorn,A., Gibson,N.J., Rivers,D.M., Hildebrand,J.G. and Morton,D.B.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1998) ARL Division of Neurobiology, University of

Arizona, Gould-Simpson Bldg. Rm 611, Tucson, AZ 85721-0077, USA

FEATURES

Location/Qualifiers
1. .3691
/organism="Manduca sexta"
/db_xref="taxon:7130"
168. .1970
/note="GC-bl"
/codon_start=1
/product="soluble quanylyl cyclase beta-1 subunit"
/protein_id="AAC61264.1"
/db_xref="GI:3372756"
/translation="MTGFVYALELLVMTFDEEWETIKKKADVAMEGSLVRLVRIE
DEITNLTAAVEVLIQIPADAILELFGKTFEFCQDSYDKILQVLGATPRDFLQNL
GLHDLGTLIPGSRSPEDTERPDLALVHYSDRPGLEHIVIGIVKTVASKLHNT
EYKVELTKKECDHVOFLITSTTVGRVSAPEIARIETLSLEPKVSPATFCRVERPF
KMDRLNTOVQAGTASRLRPVTPGCKITDVLDPVPHLEMTFANVLAHLINTVYVL
LTKPEMSYDPEEELASRLQOMLYIPETDVVVPQCPVSTVNDLDRGLCLADI
PHDARDLVMSFEADYKLTQNLVLELQOTFRELELEKORTDRLLYSVLPI
VATERHRPVPARYDVDTLLFSGIVGFANYCARNSDRHKAMKIVRLMDLNTYAFDV
LTPDKRNPVYKVTYGMVAVSGLPEYEVAAKHISLLALDMDLDSQTVTVDEGPV
GITIGHSGEVTVGIGHRMPRYCLFGNTVNLTSRCETTVGPCTINVSDTYNLMRE
DNHDEQFELTYGRGHVTMKGABPMOTWELTRKH"

BASE COUNT 1105 a 726 c 742 g 1118 t
ORIGIN

Query Match 22.9%; Score 559.6; DB 3; Length 3691;

Best Local Similarity 58.1%; Pred. No. 1.3e-119;

Matches 1060; Conservative 0; Mismatches 739; Indels 27; Gaps 3;

QY 85 CACCATGTACGGATTGTGAATCAGCCCTGGAGTGTGGTGTGATCCGCAATACGGCCC 144

DB 164 CAAATGTACGGTGTGTGAATATGCTCCCTCGAGTCTCTGTGATGAAAACACTTTTGACGA 223

QY 145 CGAGGTGTGGGAGACATCAAAAAGAGGCACAGTTAGATGAAGAGGACAGTTTCTGT 204

DB 224 AGAAACATGGGACGATTAAGAAAAGCGCGTGGCTATGGAAGGACAGTTTCTAGT 283

QY 205 CAGATAAATATATGATGACTCCAAACTTATGATTTGGTGTGCTGCGCAAGCAAGTCT 264

DB 284 GAGACAGATTATGAGGATGAGATCACTTATAATTTGATCAGCGCTGCCGTGAGTCT 343

QY 265 CAATCTCAATGTGGAGAAATCTCCAAATGTTGGGAAGATGTTTTCGTCTTTTGCCA 324

DB 344 GCMAATACCGCTGAGCGCTATATGAGCTATTTGGAAAACGTTCTCGAGTTTGCCA 403

QY 325 AGAATCTGTTATGATACAACTCTTGCTGCTCTGCTCTTAATGTGAGAAATTTCTACA 384

DB 404 AGATTCGGGTATGATAAATCTGCAAGTCTTGAGCCACGCCAGAGATTTCTACA 463

QY 385 GRACCTTGATGCTGCGACGACACCTTGCTACCACTACCCAGGAATCGGTGACCTTC 444

DB 464 GAATCTGGAGCGGCTGCACGATCATCTCGGGACATTTATATCTGGCATGAGATCCGCC 523

QY 445 CTTTAGTCTGCTGATGAGAAAAGGCAAGGACACTATTTTGCACCTACTACTCAGAGAG 504

DB 524 GTTCCGATCCAGCGCTCGAAGATGGCGCCCTTGTGCTCATTTACTTCGACAGAG 583

QY 505 AGAAGGACTTCAGGATTTGTCTATTGGAATCATCAAAACAGTGGCAACAATCCATCG 564

DB 584 GCCTGCTGAGCATATCGTTATTGGAAATTTGAAGACTGTGGCAGCAAACTCCATTA 643

QY 565 CACTGAATATGACATGAAGGTTATTACGCAAGAAATGAAGAAATGTGATCATACTCAAT 624

DB 644 CACCGA---AGTGAAGTGGAAATCTTAAAGACCAAGAGGAATGTGACACGTCAGTT 700

QY 625 TTTAATTGAGAAAAGAGTCAAAAGAGAGATTTTATGAGATCTTGACAGATTTGA 684

DB 701 CTTGATCACTGAACCTTCACCACCGGTAGGTGTCTGCACCTGAGATGAGAGAGATT-- 758

QY 685 AGAAATGTACCCAGGAATACGATCAGCCCATATACATTTCTGCAAGCTTTTCCCTTT 744

DB 759 -GAAAGTTGTCTCTTGACCAAAAGTGATCCAGGACATTTCTGCCGAGTGTTCATTT 817

QY 745 TCATATAAATATTGACCGGACCTAGTGTCTACTCAGTGTGGCAATGCTATATACAGAT 804
DB 818 CCATTTGATGTTTGTATCGTGTGAACATCGTCCAAAGCGGACGTAATGCTCTGCT 877
QY 805 TCTCCCCACGCTCCAGCCTGGAAATGAGCCCTTCTGTCTGTCTCTCCGTTGCTGCC 864
DB 878 ACTCCCAAGAGTTACCAGACACAGGCTGTAAAGATTACTGATGCTCTTGGAATGCT 937
QY 865 TCATATATATATTAGTTTCCATGGGATCCTTTCTACATCAATACATGTTTCTATTGAG 924
DB 938 CCACCTTAGAGATGACGTTTGTCTTAACTCTGCTCATATTAACACGGTCTACGTTTGA 997
QY 925 AAGCAAGGAAGATTGTGGATCTGGAGAAATAGAAATGTGAGGATGAATGACTGGGAC 984
DB 998 AACCAACCTGAAGAGATGAGTCTACTGATC-----CTCATGA 1036
QY 985 TGAGATGACCTGCTTACGCTCAAGGTCAAATGATCTACTTACCTGGAAGACAGATAGCAT 1044
DB 1037 GGAATATCGCTTCTTCTGCTTGAAGGTCAAATGCTGTACATTCAGAAACAGACGTGCT 1096
QY 1045 ACTTTTCTATGCTTACCAAGTGTCTAAGAACTGGACGATTTGACAAGGAGGAGGCTGA 1104
DB 1097 TGTATTCAGTGTATCTCAAGTGTGACCAATTTGGACGATCTGCTGCCGCGCCATAG 1156
QY 1105 TCTAAGTGACATCCCTCTGATGATGCCACGCGCATCTTGTCTTTTGGGAGAACATTT 1164
DB 1157 CATTTGCTGACATTCACCTTTCAGATGCGACGACGAGACCTGTTCTCATGCTCCGAGCAAT 1216
QY 1165 TAGAGAGGAATCAAACTCACCACCAAGAACTGGAATCTTCACTGACAGGCTACAGTCA 1224
DB 1217 TGAAGCTGATTAATAAGCTGACCCAGAACTGTGAAGTTTGTGACTGACAACTTCAACAA 1276
QY 1225 GTTAAAGGCCCTTGAAGATGAAAAGAAAAGACAGACACATTTGCTGTATTTCTCTCTCC 1284
DB 1277 ATTCGAGAGCTTGGAAATTTGGAACAAAGACAGACGAGCTTCTTACTCAGTCTGCC 1336
QY 1285 TCCGCTGTGTGCAATGAGCTGCGGACAAAGCGTCAGTGCCTGCCAAAGATATGACAA 1344
DB 1337 TATAAGCGTGGCAGCTGAACTTCGTCACCGTCCCGGTACCAGCGCGCTTACGACAC 1396
QY 1345 TGTGACCATCTTCTTAGTGGCATTTGCGGCTTCAATGCTTCTGTGCAAGCATGCATC 1404
DB 1397 CGTGACCTTCTTCTCAGCGGTATTGTAGGGTTCGCAACTACTGCGGTAGGAATCTGA 1456
QY 1405 TGGAGAAGGAGCATGAGATCGTCAACCTCTCAACGACCTCTACACAGATTTTGACAC 1464
DB 1457 CCATAAGGCGCTATGAGATCGTGAAGATGCTCAATGATCTGTACACCGCTTGTGATG 1516
QY 1465 ACTGACTGATTCGCGGAAAACCCATTTGTTTATAGGTGGAGACTGTTGTTGACAAATTA 1524
DB 1517 TCTCAGCGATCTTAAGAGGAATCTTAATGTGTATAAGGTGGAGACTGTGGGTGATAATA 1576
QY 1525 TATGACAGTGTGTTTACCAGAGCATGCAATCACCATGACGATCCATCTGCCACCT 1584
DB 1577 CATGCCGTATCGGCTTACCAGAAATATGAGGTAGTCTCATGTAAACATTTTCGCTTCT 1636
QY 1585 GGCCTTGGACATGATGGAATTTGCGGCCAGGTTTCAAGTAGATGTTGATCTGTTCAAT 1644
DB 1637 TGCTCTGGACATGATGGACTTCTCCAAACCGTGACTGTGGATGGAGAACAGTGGCAT 1696
QY 1645 AACATAGGGATACACATGGAGAGTGTACAGGTGTCTAGGACAGCGGATGCTCG 1704
DB 1697 AACCATCGGAATACACTCAGGCGAAGTGTGACTGTGTCTATCGTCCGCTACCGCATG 1756
QY 1705 ATACTGTCTTTTGGGAATACTGTCAACCTCAACGCGCAACAGAAACACAGGAGAAA 1764
DB 1757 CTACTGTCTCTCGGCAATACAGTGAATTTGACCGCGCTGTGAGACTTCTGGAGTACC 1816
QY 1765 GGGAAAAATAATGCTCTGTAATATACATACAGATGCTTATGTCTCCAAAAATTCAGA 1824
DB 1817 TGGCACCATTAACTGAGGAGGACATTAACATTTACCTTTATGCGGAGAGACACCA 1876
QY 1825 TCCCAATTCACCTTTGGAGACAGAGGCCAGTGTCCATGAAAGGGGCAAAAAAAGAACCAAT 1884

```

Db      1877 CGACAGTTGAGCTGACATCGTGGTCACCTGACCATGAGGGGAGGCGAGAACCCCAT 1936
      1885 GCAAGTTGGTTCTTATCCAGAAAAA 1910
      1937 GCAGACCTGGTTCTTCCCTCACCAGGAAGA 1962

RESULT 12
AC114761
LOCUS      70709 bp      DNA      linear      PRI 12-JUN-2002
DEFINITION Homo sapiens BAC clone RP11-354G23 from 4, complete sequence.
ACCESSION  AC114761
VERSION    AC114761.3 GI:21307587
KEYWORDS   HTG.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 70709)
            Sulston, J.E. and Waterston, R.
            Toward a complete human genome sequence
            Genome Res. 8 (11), 1097-1108 (1998)
            99063792
            PUBMED
            9847074
            2 (bases 1 to 70709)
            Buatsi, D., Bielicki, L. and Creason, K.
            The sequence of Homo sapiens BAC clone RP11-354G23
            Unpublished (2001)
            3 (bases 1 to 70709)
            Waterston, R.H.
            Direct Submission
            Submitted (11-MAR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            4 (bases 1 to 70709)
            Waterston, R.H.
            Direct Submission
            Submitted (30-APR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            5 (bases 1 to 70709)
            Waterston, R.H.
            Direct Submission
            Submitted (01-JUN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            6 (bases 1 to 70709)
            Waterston, R.
            Direct Submission
            Submitted (12-JUN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Jun 1, 2002 this sequence version replaced gi:20340523.
            -----
            Center: Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            -----
            Summary Statistics
            -----
            Center project name: H_NH0354G23
            -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-588K22, 2000 bp overlap. Actual start of this clone is at base position 88662 of RP11-588K22; actual end is at base position 70709 of RP11-354G23.

Polymorphisms have been identified between AC021433 and AC114761.

Data from AC021433 was used to finish this clone, AC114761.

FEATURES

FEATURES	Location/Qualifiers
source	1..70709
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
repeat_region	/clone="RP11-354G23"
	/clone_lib="RP11-11"
	2..101
	/rpt_family="MaLR"
repeat_region	155..590
	/rpt_family="MaLR"
repeat_region	1228..1592
	/rpt_family="MaLR"
repeat_region	1593..1726
	/rpt_family="MIR"
repeat_region	1871..2030
	/rpt_family="MIR"
repeat_region	2905..2925
	/rpt_family="(CAA)n"
repeat_region	3372..3615
	/rpt_family="MIR"
repeat_region	5097..5399
	/rpt_family="Alu"
repeat_region	6270..7028
	/rpt_family="L1"
misc_feature	8582..9872
	/note="CpG island (8GC-71.1, o/e=0.70, #CpGs=119)"
repeat_region	12310..12409
	/rpt_family="MER1_type"
repeat_region	12422..12462
	/rpt_family="MER1_type"
repeat_region	12867..12978
	/rpt_family="Alu"
repeat_region	13369..13537
	/rpt_family="MIR"
repeat_region	15390..15574
	/rpt_family="L2"
repeat_region	15920..16147
	/rpt_family="L1"
repeat_region	16537..16741
	/rpt_family="MIR"
repeat_region	18011..18038
	/rpt_family="AT-rich"
repeat_region	18645..18713
	/rpt_family="CT-rich"
repeat_region	18757..18900

```

/rpt_family-"MIR"
19796..19907
/rpt_family-"CR1"
20476..20717
/rpt_family-"L2"
22037..22066
/rpt_family-"(TTTG)n"
22070..22357
/rpt_family-"Alu"
22891..22950
/rpt_family-"L1"
23147..23412
/rpt_family-"L2"
23413..23624
/rpt_family-"Alu"
23625..24290
/rpt_family-"L2"
24309..24331
/rpt_family-"AT-rich"
24931..25231
/rpt_family-"Alu"
25653..26056
/rpt_family-"L1"
26057..26323
/rpt_family-"Alu"
26324..26432
/rpt_family-"L1"
26557..26906
/rpt_family-"MER2_type"
28704..28787
/rpt_family-"MER1_type"
28788..28926
/rpt_family-"Alu"
28927..28966
/rpt_family-"(CATA)n"
28968..29116
/rpt_family-"MER1_type"
29512..29965
/rpt_family-"L2"
31294..31344
/rpt_family-"AT-rich"
31712..32026
/rpt_family-"Alu"
32194..32590
/rpt_family-"MaLR"
33530..33559
/rpt_family-"AT-rich"
35905..36095
/rpt_family-"Alu"
36096..36410
/rpt_family-"Alu"
36411..36544
/rpt_family-"Alu"
37678..38283
/rpt_family-"MaLR"
38324..38464
/rpt_family-"MIR"
40010..40030
/rpt_family-"(TTTTTG)n"
40569..40613
/rpt_family-"Mariner"
40758..41052
/rpt_family-"Alu"
41089..41398

```

Query Match 21.3% Score 521; DB 9; Length 70709;
 Best Local Similarity 100.0%; Pred. No. 1.1e-110; Mismatches 0; Indels 0; Gaps 0;
 Matches 521; Conservative 0

QY 1923 AGGAACAAGCAGGATGACTGATCTTGGATTATGGGTGAAGAGGATACAGACT 1982
 |||||

DB 56022 AGGAACAAGCAGGATGACTGATCTTGGATTATGGGTGAAGAGGATACAGACT 56081
 |||||

```

QY 1983 AGTTCCAGTTTCTCTACACGTGCCAGCCAGGAGCAGTCTTCCCTATGGATACA 2042
|||||
DB 56082 AGTTCCAGTTTCTCTACACGTGCCAGCCAGGAGCAGTCTTCCCTATGGATACA 56141
|||||
QY 2043 GATTTCTTTTGTGCTTGTCCATACCCCAAGACTTTCTTCTAGATATATCTCTACTAT 2102
|||||
DB 56142 GATTTCTTTTGTGCTTGTCCATACCCCAAGACTTTCTTCTAGATATATCTCTACTAT 56201
|||||
QY 2103 CCGTTATTCACCTTACCTCTGCTTCTTCTATTTAGGCTTTAGTATATCTCTAAA 2162
|||||
DB 56202 CCGTTATTCACCTTACCTCTGCTTCTTCTATTTAGGCTTTAGTATATCTCTAAA 56261
|||||
QY 2163 GTTTGGCTTTTGATGTGGATGATGTGCTTCTATGTTCTTAAATCTACTACAAGCAT 2222
|||||
DB 56262 GTTTGGCTTTTGATGTGGATGATGTGCTTCTATGTTCTTAAATCTACTACAAGCAT 56321
|||||
QY 2223 ACCTAACATGGTATCTGCAAGTAGGACCCCAATATTTTGTGAATTTAGTTAA 2282
|||||
DB 56322 ACCTAACATGGTATCTGCAAGTAGGACCCCAATATTTTGTGAATTTAGTTAA 56381
|||||
QY 2283 ATGAAACTGAACAGTGTGTTGGCCATGTGTATATTTATATCATCTTTTACCAATCTGTTTA 2342
|||||
DB 56382 ATGAAACTGAACAGTGTGTTGGCCATGTGTATATTTATATCATCTTTTACCAATCTGTTTA 56441
|||||
QY 2343 GTGTTCCACATATATATATATATTTTAAATGACTATAATTAATAAGTTTATATCA 2402
|||||
DB 56442 GTGTTCCACATATATATATATATTTTAAATGACTATAATTAATAAGTTTATATCA 56501
|||||
QY 2403 TGTGTTGTATATATATATATATTTTAAATGACTATAATTTTAAAGGAGT 2443
|||||
DB 56502 TGTGTTGTATATATATATATATTTTAAATGACTATAATTTTAAAGGAGT 56542
|||||

```

RESULT 13

AC021433

LOCUS

DEFINITION

AC021433

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Bozulavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
 Campopiano, A., Castie, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hags, B., Heaford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., Lacombe, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE JOURNAL

COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 5, 2000 this sequence version replaced gi:6778505.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5027

Center clone name: 688_K_17

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 170986 bases at least Q40

Consensus quality: 174287 bases at least Q30

Consensus quality: 175619 bases at least Q20

Insert size: 174000; agarose-fp

Insert size: 176456; sum-of-contigs

Quality coverage: 5.1 in Q20 bases; agarose-fp

Quality coverage: 5.0 in Q20 bases; sum-of-contigs.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 2292: contig of 2292 bp in length
 * 2293 2392: gap of 100 bp
 * 2393 4948: contig of 2556 bp in length
 * 4949 5048: gap of 100 bp
 * 5049 8358: contig of 3310 bp in length
 * 8359 8458: gap of 100 bp
 * 8459 12531: contig of 4073 bp in length
 * 12532 12631: gap of 100 bp
 * 12632 17861: contig of 5230 bp in length
 * 17862 17961: gap of 100 bp
 * 17962 28411: contig of 10450 bp in length
 * 28412 28511: gap of 100 bp
 * 28512 40270: contig of 11759 bp in length
 * 40271 40370: gap of 100 bp
 * 40371 56792: contig of 16422 bp in length
 * 56793 56892: gap of 100 bp
 * 56893 80742: contig of 23850 bp in length
 * 80743 80842: gap of 100 bp
 * 80843 111092: contig of 30250 bp in length

* 111093 111192: gap of 100 bp
 * 111193 144467: contig of 33275 bp in length
 * 144468 144567: gap of 100 bp
 * 144568 177556: contig of 32989 bp in length.

FEATURES source

1..177556
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /map="4"
 /clone-lib="RPC1-11 Human Male BAC"

1..2292

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

34545 35422: contig of 878 bp in length
35423 gap of unknown length
36294: contig of 872 bp in length
36295 gap of unknown length
37161: contig of 867 bp in length
37162 gap of unknown length
38046: contig of 885 bp in length
38047 gap of unknown length
38943: contig of 897 bp in length
38944 gap of unknown length
39821: contig of 878 bp in length
39822 gap of unknown length
40724: contig of 903 bp in length
40725 gap of unknown length
41524: contig of 800 bp in length
41525 gap of unknown length
42428: contig of 904 bp in length
42429 gap of unknown length
43297: contig of 869 bp in length
43298 gap of unknown length
44181: contig of 884 bp in length
44182 gap of unknown length
45034: contig of 853 bp in length
45035 gap of unknown length
45922: contig of 888 bp in length
45923 gap of unknown length
46818: contig of 896 bp in length
46819 gap of unknown length
47715: contig of 897 bp in length
47716 gap of unknown length
48609: contig of 869 bp in length
48609 gap of unknown length
49478: contig of 904 bp in length
49478 gap of unknown length
50382: contig of 870 bp in length
50382 gap of unknown length
51251: contig of 870 bp in length
51252 gap of unknown length
52145: contig of 894 bp in length
52146 gap of unknown length
53042: contig of 897 bp in length
53043 gap of unknown length
53940: contig of 898 bp in length
53941 gap of unknown length
54805: contig of 865 bp in length
54806 gap of unknown length
55676: contig of 871 bp in length
55677 gap of unknown length
56569: contig of 893 bp in length
56570 gap of unknown length
57455: contig of 886 bp in length
57456 gap of unknown length
58312: contig of 857 bp in length
58313 gap of unknown length
59194: contig of 882 bp in length
59195 gap of unknown length
60086: contig of 892 bp in length
60087 gap of unknown length
60940: contig of 854 bp in length
60941 gap of unknown length
61835: contig of 895 bp in length
61836 gap of unknown length
62731: contig of 896 bp in length
62732 gap of unknown length
63619: contig of 888 bp in length
63620 gap of unknown length
64445: contig of 826 bp in length

Query Match 18.5%; Score 452.2; DB 2; Length 143255;
Best Local Similarity 95.7%; Pred. NO. 1.2e-94;
Matches 463; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
1960 TGGGGTGAAGAGGAGTACAGACTAGTTCCTCCAGTTTCTCTTAACAGGTGCCAAGCCAGG 2019

Db 20378 TNGAAGAGGAGTACAGNACTAGGGTTCCAGTTTACTCTTAACAGGTGCCAAGCCAGG 20437
QY 2020 AGCAGTTCTTCCCTATGGATACAGATTTCTTTTGTCCCTTGTCCATTATCCCAAGACTTT 2079
Db 20438 AGCAGTTCTTCCCTATGGATACAGATTTCTTTTGTCCCTTGTCCATTATCCCAAGACTTT 20497
QY 2080 CTTCTAGATATATCTCTCACTATCCGTTTATCAACCTTATAGCTCTGCTTTCTTACTTTT 2139
Db 20498 CTTCTAGATATATCTCTCACTATCCGTTTATCAACCTTATAGCTCTGCTTTCTTACTTTT 20557
QY 2140 TAGGCTTTAGTATATATCTCTCACTATCCGTTTATCAACCTTATAGCTCTGCTTTCTTACTTTT 2199
Db 20558 TAGGCTTTAGTATATATCTCTCACTATCCGTTTATCAACCTTATAGCTCTGCTTTCTTACTTTT 20617
QY 2200 TCTTAAATCTACTACAAGCATTACCTAATCAATGGTGTATCTGCAAGTAGTAGGCACCAAT 2259
Db 20618 TCTTAAATCTACTACAAGCATTACCTAATCAATGGTGTATCTGCAAGTAGTAGGCACCAAT 20677
QY 2260 AAATATTTGTTGAATTTAGTTAAATGAAGTGAACAGTGTGGGCAATGTATATTTAT 2319
Db 20678 AAATATTTGTTGAATTTAGTTAAATGAAGTGAACAGTGTGGGCAATGTATATTTAT 20737
QY 2320 ATCATGTTTACCAATCTGTTTAGTGTCCACATATATGATATATATTTAAATGACT 2379
Db 20738 ATCATGTTTACCAATCTGTTTAGTGTCCACATATATGATATATTTAAATGACT 20797
QY 2380 ATAATGTAATAAAGTTTATATCATGTTGGTGTATATCATATAGAATCATTTCTAAAG 2439
Db 20798 ATAATGTAATAAAGTTTATATCATGTTGGTGTATATCATATAGAATCATTTCTAAAG 20857
QY 2440 GAGT 2443
Db 20858 GAGT 20861

RESULT 15
DMU27123
LOCUS 2840 bp mRNA linear INV 17-FEB-1996
DEFINITION Drosophila melanogaster soluble guanylyl cyclase beta subunit (dgcbl), complete cds.
ACCESSION U27123.1 GI:861202
KEYWORDS Drosophila melanogaster.
SOURCE Drosophila melanogaster.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2840)
AUTHORS Shah,S. and Hyde,D.R.
TITLE Two Drosophila genes that encode the alpha and beta subunits of the brain soluble guanylyl cyclase
J. Biol. Chem. 270 (25), 15368-15376 (1995)
MEDLINE 95318108
PUBMED 7797526
REFERENCE 2 (bases 1 to 2840)
AUTHORS Shah,S. and Hyde,D.R.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1995) David R. Hyde, Biological Sciences, Univ. of Notre Dame, Galvin Life Science Building, Notre Dame, IN 46556, USA

FEATURES
source
1. .2840
/organism="Drosophila melanogaster"
/strain="Oregon-R"
/db_xref="taxon:7227"
/chromosome="3"
/map="100B"
1. .2840
/gene="dgcbl"
300. .2576
/gene="dgcbl"

gene
CDS

```
/codon_start=1
/product="soluble guanylyl cyclase beta subunit"
/protein_id="AA87941.1"
/db_xref="GI:861203"
/translating="MYGFVYALELVLKHFGEIWEIKKKAMVNEGOFVLYRLOYD
DEITYNLGNAVEILNIPADDILELFGKFFFCQDSGYDKILQVIGATPRDFLQNLID
ALHDHGLTLPGRAPSFRCETKDEGLLHYISERFGLHIVIGIYKAVASKLHGVVEV
EIDIVRRKGPIDAEKERAIRENQOLLEDAVATTTGTSATVYLAPSSDAERNNNH
GSGSNNGMANNNTVNNVNDGQIASSETDPSIALSTCPIDAFSDFCDDGKQK
LRLKKNKDDIERVDFVFLIREINAAKSOVDKDEVPDDMEFLCEAPLISPATFC
KVPFHLMPDROMKIYQAGKAVSRVIRVAEENCLEVEEARPHQLQNFENLISHI
NIYIYQTRGAMSSHEQREFRLKGMWYIPETDRILLFQCYFSVNNDDLTKKGLYI
SDVPLHDAADLVLLSEKFEAEYKLTNLEMLFDKLOQTFRDLESCKQKTDRLLYSVL
PKSANELRHQRPVPRKRVDSVTFMFSGVIGFQCYCAANTDPDGANKIYKMLNELTYV
FDALTSKRLNLYKVTVDGKYMYVSLPDHCEHAKMARVALDMDMAMKVKWNGS
NPQIYITIGHSQVWTVGIGNRPRYCLFGNTVNLSTRTEITGVPGRINVSSEYRLL
CMAINDODSPHLEVRGPIVIMKGRPTPMDCHFLTRATVSSILGGTSS7SGSGGGGNGSLD
SPLLQATPTAGGPRFP"
BASE COUNT      730 a   753 c   782 g   573 t   2 others
ORIGIN
Query Match      15.2% Score 371.4; DB 3; Length 2840;
Best Local Similarity 58.1% Pred. No. 9.7e-76;
Matches 694; Conservative 0; Mismatches 471; Indels 30; Gaps 1;
QY 705 CAGCATACGCCCATATACATTCGCAAGCTTTTCCCTTTTCATATAATATTTGACGGG 764
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1276 CACTCATATCGCGCGCACTTTTCGCAAGCTTTCCCTTCCACCTGATGTTGATAGGC 1335
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 765 ACCTAGTGGTCACGTGTGGCAATGCTATATACAGAGTTCCTCCCGACCTCCAGCTG 824
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1336 AATGAAATCGTTCAGCGCGCAAGCTGTCTCGAGTAATTCGCAAGTGTGCCGAGG 1395
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 GGAATTCGACCCCTTCTCTCTCTCTCGCTGCTGCTCTCATATGATATTAGTTTCC 884
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1396 AGAAGTTCCTTAATAGAGGTGGTGGAGGCCATCGCGCGACCTGCACCTCACTTCG 1455
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 ATGGGATCTTCTCAGATCAATACACTGTTTGTATTTGAGAGCAAGGAGGATTTG 944
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1456 AGAATATCTCTCCACATAAACCACATCTATCTCTGACACACGTCAGGGGGCCATGA 1515
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 945 ATGTGGAGAAATAGATGTGAGGATCAACTGACTGGGACTGAGATCAGCTGCTTAGCTC 1004
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1516 GC-----AGCGGCGACGAGCAGCGGTTTCTGAGAC 1545
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1005 TCAAGGCTCAATGATCTACTTACCTGAAGCAGATAGCATACTTTTCTATGTTCAACAA 1064
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1546 TGAAGGGCAATGATGTACATTCGGAACGATCGGATCTGTTCAATGCTATCCCA 1605
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1065 GTGTATGAACCTGGAGGATTTGACAGGAGAGGGGTGATCTAATGATGACATCCCTCTGC 1124
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1606 GTGTATGAATCTGGATGATCAACCAAGAGGAGCTTACATCTCGGATGTGCGGTGC 1665
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1125 ATGATCGACGCGGATCTGTTCTTTTGGAGACAAATTTAGAGAGGATACAAACTCA 1184
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1666 ACAGCGTCCGAGGATTTGGTCTGCTCTCCGAGAGTTTCGAGCGGAGTCAAGCTCA 1725
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1185 CCCAAGACTGGAATCCTCACTGACAGGCTACAGCTCAGTTTAAGAGCCCTCGAAGATG 1244
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1726 CCAAGAACCTGGAATGCTCACCGCAAGCTGCAGCAGACATTCGCGGATCTCGAGAGCG 1785
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1245 AAAAGAAAAGACAGACATTTGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1304
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1786 AGAAGCAGAGAGACCGAGCGCTGCTCTATTCGGTGTCTGCAAAAGTCGGTGGCCAAATGAGT 1845
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1305 TCGCGGACAGGCTCCAGTCCCTGCAAAAGATATGACATGTGACCATCTCTCTCTCTCTCT 1364
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1846 TAGCAGATCAGCGCTCTGTCGCCCCCAACAGCTACGCTCCGTCGAGCTTGATGTTCTCCG 1905
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1365 GCATTGTGGGCTTCAATGCTTTCTGTAGCAAGCATGCTATCTGGAGAGGAGCGCATGAAGA 1424
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1906 GCATCGTGGGCTTTGGCCAAATACTGTGCGGCCAACACAGATCCCGAGCGGGCAATGAAA 1965
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 1, 2003, 16:02:36

Job time : 4194.4 secs